(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 February 2001 (08.02.2001)

PCT

(10) International Publication Number
WO 01/09300 A2

(51) International Patent Classification[7]: C12N 15/00

(21) International Application Number: PCT/NL00/00534

(22) International Filing Date: 27 July 2000 (27.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
99202540.3 2 August 1999 (02.08.1999) EP

(71) Applicant (*for all designated States except US*): **KEYGENE N.V.** [NL/NL]; P.O. Box 216, NL-6700 AE Wageningen (NL).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **FIERENS-ONSTENK, Bernarda, Gerharda, Johanna** [NL/NL]; Vrekenhorst 91, NL-3905 VS Veenendaal (NL). **DE BOTH, Michiel, Theodoor, Jan** [NL/NL]; Gruttoweide 199, NL-6708 BG Wageningen (NL).

(74) Agent: **JORRITSMA, Ruurd**; Nederlandsch Octrooibureau, Scheveningseweg 82, P.O. Box 29720, NL-2502 LS The Hague (NL).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** METHOD FOR GENERATING RESISTANCE AGAINST CGMMV IN PLANTS, GENETIC CONSTRUCTS FOR USE IN SAID METHOD, AND CGMMV-RESISTANT PLANTS OBTAINED VIA SAID METHOD

(57) **Abstract:** The present invention relates to a method for generating resistance against Cucumber Green Mottle Mosaic Virus (CGMMV) in plants, in particular in plants susceptible to infection by CGMMV, such as *Cucurbitaceae* species, including melon, cucumber, watermelon and bottlegourd. In the invention, the plants are provided with replicase-mediated resistance against CGMMV. For this purpose, the plants are transformed with a polynucleotide sequence that: i) has been derived from the 129 kD sequence, the 57 kD sequence, or the 186 kD readthrough sequence of native CGMMV; ii) upon (at least) transformation into the plant and transcription into RNA - and usually also translation into the corresponding encoded protein - can provide the plant with resistance against CGMMV; but iii) does not encode any replicase activity. In particular, this will be a nucleotide sequence that encodes a defective variant of the replicase of CGMMV. The invention further relates to genetic constructs containing such a polynucleotide sequence suitable for use in the above method, and to CGMMV-resistant transgenic plants obtained via said method.

WO 01/09300 A2

Method for generating resistance against CGMMV in plants, genetic constructs for use in said method, and CGMMV-resistant plants obtained via said method.

The present invention relates to a method for generating resistance against Cucumber Green Mottle Mosaic Virus (CGMMV) in plants, in particular in plants that are susceptible to infection by CGMMV, such as species of the *Cucurbitaceae* family.

The invention further relates to genetic constructs suitable for use in said method, and to CGMMV-resistant transgenic plants obtained via said method.

Methods of introducing DNA sequences into the genome of plants have been known for many years and have been widely used to alter the properties of plants varieties. Such methods are among others *Agrobacterium*-mediated transformation (Horsch *et al.*, 1985; Rogers *et al.*, 1986), protoplast transformation using electroporation or other techniques to introduce naked DNA molecules into the plant cell (Shillito *et al.* 1985), and particle bombardment to introduce naked DNA molecules into plant cells or tissues (Christou *et al.*, 1994).

Among the most important applications of plant genetic engineering are those aimed at introducing resistance genes to a wide variety of plant pests and plant pathogens, such as bacteria, fungi, nematodes, insects and viruses. Many examples of virus resistance in a wide variety of plant species have been described over the last decades (Wilson *et al.*, 1993). The various methods to obtain virus resistance in plants through the introduction of gene sequences are either based on the use of genes of plant origin; on the use of sequences/genes derived from the viral pathogen itself (so-called pathogen-derived resistance (Wilson *et al.*, 1993), or on the use of genes of yet different origin. Sequences originating from the viral genome can be either cloned or PCR-amplified DNA sequences obtained from the genome of DNA viruses, such as geminiviruses (Kunik *et al.*, 1994) or the cDNA sequences obtained from the genomes of RNA viruses through the use of cDNA cloning or RT-PCR amplification.

Examples of sequences/genes of RNA viruses that have been succesfully used in the engineering of virus resistance in plants include:

1. coat protein genes of tobamoviruses, cucumoviruses, potyviruses, potexviruses (Beachy *et al.*, 1990);
2. RNA-dependent RNA polymerase genes (replicase genes) of tobamoviruses, cucumoviruses, potyviruses (Anderson *et al.*, 1992; Donson *et al.*, 1993; Audy *et al.*, 1994);

3. nucleoprotein genes of tospoviruses (Goldbach and De Haan, 1993; Prins *et al.*, 1994; Vaira *et al.*, 1995);
4. movement protein genes of tobamoviruses and cucumoviruses (Cooper *et al.*, 1995).

Cucumber Green Mottle Mosaic Virus (CGMMV) is a member of the tobamovirus group and infects plant species of the *Cucurbitaceae* family: melon (*Cucumis melo*), cucumber (*C.sativus*), watermelon (*Citrullus vulgaris*) and bottlegourd (*Lagenaria siceraria*), but not apparently *Cucurbita pepo* (squash, pumpkin, courgette). The host range of the virus is basically restricted to members of the *Cucurbitaceae*, with the exception that some but not all strains also infect tobacco and/or the diagnostic species *Datura stramonium* and *Chenopodium amaranticolor* (Hollings *et al.*, 1975).

Several different strains can be distinguished serologically and by their response in *C.amaranticolor* and *D. stramonium* (Hollings *et al.*, 1975) The "*type strain*" was originally identified in Europe and does not normally cause fruit symptoms in cucumber. Another European strain, called the *cucumber aucuba mosaic* strain, *cucumber virus 4* or *Cucumis virus 2A* causes fruit symptoms in cucumber. A number of strains are known from Japan. In watermelon, the *watermelon strain* causes serious disease, whereas the *Japanese cucumber strain* (also called Kyuri Green Mottle Mosaic Virus) and the *Yodo strain* cause fruit distortions in cucumber. The *CGMMV-C* strain from India is a pathogen on bottlegourd and serious infections can cause complete crop losses.

In cucumber, CGMMV causes vein clearing, a light and dark green leaf mottle, leaf blistering and malformation and stunted growth, seriously affecting fruit yield. The East European isolates of the *aucuba mosaic strain* produces bright yellow leaf mottling and fruit discolouration.

CGMMV is transmitted through seed, but mostly through mechanical infection via the roots in contaminated soil, and through foliage contact and handling of plants (Hollings *et al.*, 1975). The virus particles are extremely stable and survive several months at normal temperatures. This stability combined with the very high infectivity through mechanical contact of the foliage is responsible for the economic importance of this virus, as even one or a few infected plants in a cucumber greenhouse can eventually cause the infection and loss of the total crop. Also, infection may not only spread rapidly over a current crop, but also – due to the strong persistance of the virus – affect subsequent crops. Therefore, a CGMMV infection may require sterilisation of an entire greenhouse, as well as the use of sterile tools and materials.

The complete sequence of only one isolate of CGMMV has been determined (Ugaki *et al.*, 1991; Genbank accession numbers D12505 and D01188). This isolate "SH" had been found in infected watermelon plants in East Asia. Furthermore, the sequence of the coat protein gene of one other isolate ("W") obtained from infected watermelon is known (Meshi *et al.*, 1983; Genbank accession numbers V01551 and J02054), as well as the sequence of the 29 kD movement protein gene of a watermelon strain (Saito *et al.*, 1988; Genbank accession number J04332). The nucleotide sequence of the CGMMV-SH isolate shows 55 to 56% identity with tobacco mosaic virus (TMV) and tobacco mild green mosaic virus (TMGMV), both other members of the tobamovirus group (Ugaki *et al.*, 1991).

As described by Ukagi *et al.*, the genome of CGMMV consists of a single-stranded RNA molecule coding for at least four open reading frames, encoding putative proteins of 186 kD, 129 kD, 29 kD and 17.3 kD, of which the 17.3 kD ORF is known to encode the coat protein. In this respect, Ugaki *et al.* state: "*No CGMMV-encoded proteins except for the coat protein have yet been identified in vivo*".

The CGMMV genome is schematically shown in Figure 1. As can be seen therein, the ORF encoding the 186 kD protein starts at the same site as the ORF encoding the 129 kD protein, and adds a putative 57 kD polypeptide to the 129 kD ORF. The presence of this 57 kD protein alone has not been detected in infected plants. Instead, the 186 kD protein has been found, being the product of a read-through translation of the 129 kD and the 57 kD ORFs.

This 186 kD protein is thought to play a role in virus replication. Also, the 129 kD ORF is thought to encode a replicase function, whereas the 29 kD ORF is thought to encode a movement protein.

Hereinbelow, the nucleotide sequence corresponding to the ORF encoding the 129 kD protein will be referred to as "*129 kD sequence*", the sequence corresponding to the 186 kD readthrough protein will be referred to as "*186 kD sequence*", and the nucleotide sequence corresponding to the ORF encoding the 57 kD readthrough part will be referred to as "*57 kD sequence*". These nucleotide sequences and the corresponding protein sequences are given in the sequence listings, as further described below.

Object of the invention was to provide a method for protecting plants, in particular plants susceptible to infection with CGMMV such as species of the *Cucurbitaceae* family, against infection with CGMMV, and in particular against infection with strains of CGMMV prevalent in Europe, such as the strains encountered in the cultivation of cucumbers in greenhouses.

Further objects were to provide means for use in said method, in particular a genetic construct that can be used for transforming plants or plant material so as to provide transgenic plants resistant against infection with CGMMV. Further objects of the invention will become clear from the description given hereinbelow.

For these purposes, applicant has investigated the symptomatology and the nucleotide sequence of the coat protein genes of 10 European strains of CGMMV, and compared these with the SH strain described by Ugaki *et al.*. A list of these strains, with their geographical origin and symptoms on cucumber, is given in Table 1.

Table 1. List of collected CGMMV-isolates with their geographical origin and symptoms on cucumber.

CGMMV isolate	**geographical origin**	**symptoms on cucumber**
1	Eastern Europe	vein clearing, mosaic
2	Eastern Europe	verin clearing, mosaic
3	IPO-DLO, the Netherlands	almost without symptoms
4	the Netherlands	weak leaf chlorosis
5	the Netherlands	weak leaf chlorosis
6	Proefstation Naaldwijk, the Netherlands	chlorosis
7	Rijk Zwaan, the Netherlands	chlorosis
8	Israël	chlorosis
9	Almería, Spain	chlorotic leaf spots
10	Almería, Spain	weak leaf chlorosis
CGMMV-SH	Japan	strong chlorotic leaf mosaic

It was found that the sequences for the 10 European isolates are highly homologous (i.e. homology on the nucleotide level of 97%), and show about 90 % homology (on the nucleotide level) with the SH-isolate. The nucleotide sequences encoding the coat proteins of each of the isolates 1-10, as well as strain SH, are given in the sequence listings, as further described below. The corresponding phytogenetic tree is shown in Figure 2. This shows that the European isolates can be considered to constitute a subgroup of the CGMMV species.

In the sequence listings:

- SEQ ID no.1 gives the nucleotide sequence encoding the 129 kD replicase protein of CGMMV isolate 4, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.2 gives the amino acid sequence of the 129 kD replicase protein of CGMMV isolate 4, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.3 gives the nucleotide sequence encoding the 57 kD protein of CGMMV isolate 4, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.4 gives the amino acid sequence of the 57 kD replicase protein of CGMMV isolate 4, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.5 gives the nucleotide sequence encoding the 186 kD readthrough protein of CGMMV isolate 4, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.6 gives the amino acid sequence of the 186 kD readthrough protein of CGMMV isolate 4, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.7 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 1, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.8 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 2, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.9 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 3, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.10 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 4, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.11 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 5, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.12 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 6, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.13 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 7, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.14 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 8, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.15 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 9, with the ORF of the coat protein starting with the ATG codon at bp 523-525;

- SEQ ID no.16 gives the nucleotide sequence encoding the coat protein of CGMMV isolate 10, with the ORF of the coat protein starting with the ATG codon at bp 523-525;
- SEQ ID no.17 gives the nucleotide sequence encoding the 129 kD replicase protein of CGMMV isolate SH;
- SEQ ID no.18 gives the amino acid sequence of the 129 kD replicase protein of CGMMV isolate SH;
- SEQ ID no.19 gives the nucleotide sequence encoding the 57 kD protein of CGMMV isolate SH;
- SEQ ID no.20 gives the amino acid sequence of the 57 kD replicase protein of CGMMV isolate SH;
- SEQ ID no.21 gives the nucleotide sequence encoding the 186 kD readthrough protein of CGMMV isolate SH;
- SEQ ID no.22 gives the amino acid sequence of the 186 kD readthrough protein of CGMMV isolate SH;
- SEQ ID no.23 gives the nucleotide sequence encoding the coat protein of CGMMV isolate SH;
- SEQ ID's nos. 24-40 give the nucleotide sequences of the primers used in the Examples;
- SEQ ID's nos. 41-44 give the nucleotide sequences used in assembling the leader sequences used in the constructs described in the Examples;

In the above sequence listings, the nucleotide sequences given are DNA sequences, as the genetic constructs of the invention described below will usually contain or consist of a DNA. As CGMMV is an RNA virus, it will be clear to the skilled person that these sequences will occur in the virus as the corresponding RNA sequence (i.e. with U replacing T). Also, it will be clear to the skilled person that the nucleotide sequences given above may be followed – both in the virus as well as in a construct of the invention – with a suitable termination codon, i.e. TAA/UAA, TAG/UAG or TGA/UGA (not shown).

Furthermore, as will be clear to the skilled person, the nucleotide sequence encoding the coat protein will usually start with an ATG codon. For example, in SEQ ID NOs 1 – 16, the nucleotide sequence encoding the coat protein starts at the ATG codon at base positions 523-525. (In the nucleotide sequence of SEQ ID NOs 1 – 16, the nucleotide sequence encoding the coat protein is preceded by another nucleotide sequence, e.g. encoding a movement protein. Accordingly, when hereinbelow reference is made to any nucleotide sequence of SEQ ID NOs 1 – 1 6, this also explicitly includes the nucleotide sequence starting at the ATG codon at base positions 523-525 of these SEQ ID's).

A particular purpose of the invention is therefore to provide a method that can provide plants with resistance against all these strains simultaneously, and more in particular a type of resistance that is agronomically useful, i.e. that can be used to protect (crops of) plants that are cultivated under circumstances wherein the high infectivity and persistance of CGMMV can be a major problem, such as the cultivation of cucumbers in greenhouses.

In the invention. this problem is generally solved by transforming a plant with a polynucleotide sequence (e.g. as part of a genetic construct) that can provide the plant with so-called "replicase-mediated" resistance against CGMMV. In particular, this wil be a polynucleotide sequence that

i) has been derived from the 129 kD sequence, the 57 kD sequence, or the 186 kD readthrough sequence of native CGMMV;

ii) upon (at least) transformation into the plant and transcription into RNA - and usually also translation into the corresponding encoded protein - can provide the plant with resistance against CGMMV; but

iii) does not encode any replicase activity.

In particular, such a polynucleotide sequence can encode a polypeptide or protein that is capable of providing a plant with resistance against GCMMV, but that by itself has no replicase activity, for instance due to one or more alterations in its amino acid sequence, compared to the amino acid sequence encoded by the 129 kD sequence, 57 kD sequence, and/or 186 kD readthrough sequence of native CGMMV.

However, according to one specific embodiment of the invention, the polynucleotide sequence may also comprise, or even consist of, the native 57 kD sequence.

In a first aspect, the invention therefore relates to a method for generating resistance in a plant against CGMMV, said method comprising at least the step of transforming said plant with a polynucleotide encoding a defective variant of the replicase gene of CGMMV.

In a further aspect, the invention also relates to a method for providing a transgenic plant and/or plant cell that is resistant against infection with CGMMV, comprising at least the step of transforming said plant or plant cell with a polynucleotide sequence encoding a defective variant of the replicase gene of CGMMV.

In another aspect, the invention also relates to a genetic construct suitable for transforming a plant, said construct at least comprising a polynucleotide sequence encoding a defective variant of the replicase gene of CGMMV, and optionally further elements of genetic constructs known per se. The invention also relates to a plant, plant cell and/or plant material that has been transformed with a genetic construct of the invention.

The invention also relates to transgenic plants that contain a polynucleotide sequence encoding a defective variant of the replicase gene of CGMMV, and/or that have been provided with resistance against infection with CGMMV by the method of the invention.

In the context of the invention, by the "*replicase gene of CGMMV*" is meant the native 129 kD sequence, the native 57 kD sequence, and/or the combined native 186 kD "readthrough" product of the native 129 kD and native 57 kD sequences.

By a "*native*" sequence is meant any RNA sequence that naturally occurs in CGMMV, including all isolates and strains thereof, as well as any DNA sequence that corresponds to these naturally occuring RNA sequences. Examples of such native sequences are the 129 kD nucleotide sequences given in SEQ. ID no.1 and SEQ. ID no.17, the 57 kD nucleotide sequences given in SEQ. ID no.3 and SEQ. ID no.19 and the 186 kD nucleotide sequences given in SEQ. ID no.5 and SEQ. ID no.21. It will be clear to the skilled person that there may be (further) naturally occuring variants of the RNA sequence from which the DNA sequences in the sequence listings were derived, and these (and the DNA sequences corresponding thereto) are also included within the term "*native* sequence".

By "*a polynucleotide sequence encoding a defective variant of the replicase gene of CGMMV*" in its broadest sense is meant a polynucleotide sequence that

i) upon (at least) transformation into a plant and transcription into RNA generates resistance against infection with CGMMV in said plant; and

ii) upon (at least) transformation into a plant and transcription into RNA does not lead to generation of (any) replicase activity in said plant (or at least -when it does lead to expression of some replicase activity - leads to expression of a replicase activity that is severely reduced compared to expression of the native gene encoding CGMMV replicase).

Herein, the terms "*plant*", "*transformed plant*" and/or "*transgenic plant*" include all parts or tissues of such a plant, including but not limited to individual cells of such a plant. These terms also includes material of or for such a plant, such as material that can be regenerated into a (mature) plant, including but not limited to protoplasts and/or callus tissue, or material that can be cultivated into a mature plant, such as cultivation material.

The plant is preferably a plant that is susceptible to infection with CGMMV, more preferably a plant belonging to the *Cucurbitaceae* family, such as melon (*Cucumis melo*), cucumber (*C.sativus*), watermelon (*Citrullus vulgaris*) and bottlegourd (*Lagenaria siceraria*).

Included within the term "CGMMV" are all known strains thereof, including those prevalent in Europe and Asia. In particular, the method of the invention can be used to protect

plants against strains of GCMMV prevalent in Europe (including Israel), such as those which are a problem in the cultivation of melons and in particular cucumbers in greenhouses, although the invention is not limited thereto.

In doing so, a major advantage of the invention is that it can provide protection against several, and preferably all, (such) strains of CGMMV simultaneously. Another advantage of the invention is that it provides "absolute" protection against CGMMV, which means that - upon expression of a polynucleotide sequence encoding a defective replicase in a plant - essentially no viral particles can be detected in the transformed plant (material). The method of the invention therefore does not lead to a deferral or slowing down of the onset of symptoms, as may occur when so-called "coat protein-mediated" resistance is used. Also, the method of the invention leads to a high level of resistance, and may also have the advantage of a favorable temperature effect.

Usually, the "*nucleotide sequence encoding a defective variant of the replicase gene of CGMMV*" will be a nucleotide sequence in which - compared to a nucleotide sequence encoding the corresponding native replicase of CGMMV - one or more nucleotides have been added, replaced and/or removed. In particular, the "*nucleotide sequence encoding a defective variant of the replicase gene of CGMMV* may be a nucleotide sequence that comprises, and preferably consists of:

- a nucleotide sequence corresponding to the native 129 kD sequence in which - compared to said native sequence - one or more nucleotides have been added, replaced and/or removed;
- a nucleotide sequence corresponding to the native 186 kD sequence in which - compared to said native sequence - one or more nucleotides have been added, replaced and/or removed, e.g. in the part of the native 186 kD sequence corresponding to the 129 kD sequence, to the 57 kD sequence, or both;
- a nucleotide sequence corresponding to the native 57 kD sequence;
- a nucleotide sequence corresponding to the native 57 kD sequence in which - compared to said native nucleotide - one or more nucleotides have been added, replaced and/or removed;

such that said nucleotide sequence is capable - upon (at least) transformation into a plant and transcription into RNA – to confer to said plant resistance against infection with CGMMV, and such that said nucleotide sequence – upon (at least) transformation into a plant and transcription into RNA - is not capable of generating of (any) replicase activity in said plant.

Usually, the "*nucleotide sequence encoding the defective variant of the replicase gene of CGMMV*" will encode a protein or polypeptide, more specifically a protein or polypeptide that:

1) upon being expressed in a plant is capable of generating resistance against CGMMV in said plant; and
2) upon being expressed in a plant has no replicase activity (or -when it has some replicase activity- has severly reduced replicase activity compared to the native CGMMV replicase).

Such a protein or polypeptide will be generally referred to hereinbelow as "*defective replicase*"; and a polynucleotide sequence encoding such a protein or polypeptide will be referred to as a "*polynucleotide sequence encoding a defective replicase*".

Usually, the defective replicase will be a derivative - such as an analog, homolog, variant, mutant, part, fragment or combination of two or more such parts or fragments, etc.- of the amino acid sequence encoded by the native 129 kD sequence, the native 186 kD sequence and/or the native 57 kD sequence, in which -compared to the amino acid sequence encoded by the corresponding native sequence- one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed, leading to loss of replicase activity (or at least an inability to generate replicase activity when expressed in the plant).

In particular, the defective replicase may be a protein or polypeptide that comprises, and preferably consists of:

- an amino acid sequence corresponding to the amino acid sequence encoded by the native 129 kD sequence, in which - compared to said native sequence - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed;
- an amino acid sequence corresponding to the amino acid sequence encoded by the native 186 kD sequence, in which - compared to said native sequence - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed, leading to loss of replicase activity;
- an amino acid sequence corresponding to the amino acid sequence encoded by the native 57 kD sequence;
- an amino acid sequence corresponding to the amino acid sequence encoded by the native 57 kD sequence, in which - compared to said native sequence - one or more

amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed;

or any combination thereof, provided that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

More in particular, the defective replicase may be a protein or polypeptide that comprises, and preferably consists of:

- an amino acid sequence corresponding to a part or fragment of the amino acid sequence encoded by the native 129 kD sequence, or to a combination of two or more such parts or fragments;
- an amino acid sequence corresponding to a part or fragment of the amino acid sequence encoded by the native 186 kD sequence, or to a combination of two or more such parts or fragments; or
- an amino acid sequence corresponding to the amino acid sequence encoded by the native 57 kD sequence.

such that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

An amino acid sequence "*corresponding to a part or fragment of the amino acid sequence encoded by the native 186 kD sequence, or to a combination of two or more such parts or fragments*" may for instance comprise: i) at least one part or fragment of the amino acid sequence encoded by the native 129 kD sequence combined with at least one part or fragment of the amino acid sequence encoded by the native 57 kD sequence (which combination of parts or fragments may or may not correspond to a contiguous amino acid sequence encoded by the native 186 kD sequence); ii) at least one part or fragment of the amino acid sequence encoded by the native 129 kD sequence combined with the full amino acid sequence encoded by the native 57 kD sequence, and/or iii) at least one part or fragment of the amino acid sequence encoded by the full native 129 kD sequence combined with at least one part or fragment of the amino acid sequence encoded by the native 57 kD sequence.

It is known, however, that expression in a plant of a nucleotide sequence encoding the full 129 kD sequence of the native replicase usually does not provide resistance against infection with CGMMV, but may even - upon infection of the plant - promote or facilitate multiplication of the virus. Therefore, in one embodiment, the invention does not comprise the expression in a plant of said replicase, nor the use of a polynucleotide sequence encoding such a replicase.

Even more preferably, the defective replicase is a protein or polypeptide that consists of:

- an amino acid sequence corresponding to a part or fragment of the amino acid sequence encoded by the native 129 kD sequence, or a combination of two or more such parts or fragments; such that the resulting protein or polypeptide shows no replicase activity, but is still capable, upon expression in a plant, to generate resistance against CGMMV in said plant; or
- an amino acid sequence corresponding to the amino acid sequence encoded by the native 57 kD sequence.

Any such parts or fragments may also contain one or more further amino acid substitutions, insertions or deletions compared to the native sequence, but this is not preferred.

Most preferably, the defective replicase is a so-called "truncated replicase", i.e. an amino acid sequence corresponding to the amino acid sequence encoded either by the 129 kD sequence and/or by the 186 kD sequence, from which - compared to the native amino acid sequence - one or more amino acid residues are lacking at the carboxyl-terminus, such that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant. (In case of a truncated replicase based upon the 186 kD sequence, this usually means that the resulting protein will contain the full amino acid sequence of the 129 kD sequence, as well as part of the amino acid sequence of the 57 kD sequence (i.e. that is contiguous to the 129 kD sequence in the amino acid sequence encoded by native 186 kD sequence), with one or more amino acids lacking at the carboxy-terminus of the 57 kD part, although the invention in its broadest sense is not limited thereto).

The polynucleotide sequence that encodes such a truncated replicase may either comprise, or preferably consist of, the full native 129 kD sequence or 186 kD sequence, respectively, in which a stopcodon has been introduced at a desired site, or a polynucleotide sequence from which – compared to the full native 129 kD sequence or 186 kD sequence, respectively – one or more codons coding for the carboxy-terminal amino acid residues have been removed, i.e. starting from the 3'end of the native sequence(s).

As mentioned below, preferably a stopcodon is introduced in to the native sequence, in particular in the so-called GDD motiv or in the P-loop. Examples thereof are the polynucleotide sequences comprised in the vectors shown in Figures 3-8, and as described in the Experimental Part.

Again, any such truncated replicase may also contain one or more amino acid substitutions, insertions or deletions compared to the native sequence, but this is not preferred.

As mentioned above, (the polynucleotide sequence encoding) the defective replicase is such that - after expression in a plant or plant cell - it is still capable of generating resistance against CGMMV in said plant. Usually, this means that the defective replicase will have at least one biological function that allows the defective replicase to protect the plant against CGMMV infection, such as for example down-regulation of viral replication or interference with the replication of the wild-type CGMMV, for instance by competing with wild-type virus for the replication machinery in the plant (cell). It will be clear that in order to achieve such a biological function, the defective replicase must usually have a certain minimal level of amino acid similarity with the amino acid sequences encoded by the native 129 kD, 186 kD and/or 57 kD sequences. In so far as the defective replicase is similar to the corresponding native amino acid sequence, this may be because it contains – on the corresponding amino acid positions - the same amino acid residues as the native amino acid sequence, or amino acid residues comparable thereto. The latter will usually comprise so-called "conservative" amino acid substitutions, for instance involving replacing a given acidic or basic amino acid residue by another acidic or basic amino acid residu.

However, there will also be differences in amino acid sequence between the defective replicase and the native replicase (i.e. the 129 kD, 186 kD or 57 kD protein), such that the defective replicase will no longer provide replicase activity. The skilled person will be able to select appropriate alterations to the amino acid sequence of the native replicase. As will be clear to the skilled person, a single (amino acid or nucleotide) alteration may be sufficient, or two or more such alterations may be required, dependant upon the position and nature of the alteration(s) compared to the amino acid sequence of the native replicase.

Whether a given polynucleotide sequence encodes a defective replicase according to the invention – or at least is capable of protecting a plant against infection with CGMMV – can simply be tested by transforming a plant, plant cell or plant material with a construct containing said polynucleotide sequence, and then exposing the plant, plant cell, plant material, and/or a mature plant generated therefrom, to CGMMV under conditions such that infection may occur. It can then be easily determined whether the polynucleotide sequence/construct is capable of protecting the plant, i.e. by suitably determining the presence of the virus, or simply by the presence or absence of symptoms of CGMMV-infection.

In general, as a minimum, when the defective replicase contains any amino acid substitutions or insertions, it will have an amino acid homology (i.e. identity on corresponding

position) with the corresponding native replicase protein of at least 80%, preferably at least 90%, more preferably at least 95%, with amino acid deletions not being taken into account, and a single amino acid insertion being counted as a single alteration.

In general, as a minimum, when the defective replicase contains one or more amino acid deletions, it will usually contain at least 30%, preferably at least 50%, more preferably at least 70%, and usually 80-90%, and may even contain as much as 95-99%, of the amino acid sequence of the corresponding native replicase protein, with any amino acid insertions or substitutions not being taken into account.

A truncated replicase based upon the 129 kD sequence will usually contain at least 50%, preferably at least 70%, and may contain as much as 80-95%, of the amino acid sequence of the native replicase. A truncated replicase based upon the 186 kD sequence may contain the full 129 kD protein followed by one or more amino acids from the 57 kD sequence, and usually contains the full 129 kD sequence followed by 1-95%, preferably 5-50%, of the 57 kD sequence.

The differences in amino acid sequence mentioned above can be differences compared to any of the amino acid sequences given in SEQ ID's 2, 4, 6 and/or 18, 20, 22, and/or compared to any naturally occuring variant of these amino acid sequences. These differences are at least such that the resulting protein does not correspond to a naturally occuring/native protein (including those given in SEQ ID's 2, 4, 6 and/or 18, 20, 22).

The polynucleotide sequences used in the invention are such that they encode the above defective replicases. For this purpose, they may contain the same codons as in the corresponding positions on the native 129 kD, 186 kD and/or 57 kD sequence, or codons equivalent thereto due to the degeneracy of the genetic code.

The polynucleotide sequence encoding the defective replicase can be provided in a manner known per se, for instance starting from the known sequence of the native 129 kD, 57 kD and/or 186 kD sequences, and/or from a nucleic acid that encodes said sequences.Usually, this will involve introducing one or more deletions, substitutions and/or insertions of one or more nucleotides, or even of one or more codons into, or compared to, the native sequence. Such deletions, substitutions and/or insertions will be collectively referred to hereinbelow as "*alterations*".

Accordingly, the polynucleotide sequence encoding the defective replicase may be a sequence that contains one or more such alterations compared to any of the nucleotide sequences given in SEQ ID's 1, 3, 5 and/or 17, 19, 21, and/or compared to any naturally occuring variant of these nucleotide sequences (including DNA sequences corresponding to

the RNA sequences as present in the virus). These differences are at least such that the protein encoded by the polynucleotide sequence does not correspond to a naturally occuring/native protein (including those given in SEQ ID's 2, 4, 6 and/or 16, 18, 22).

Furthermore, besides the alterations mentioned above, and compared to nucleotide sequences given in SEQ ID's 1, 3, 5 and/or 17, 19, 21 and/or compared to any naturally occuring variant of these nucleic acid sequences (including DNA sequences corresponding to the RNA sequences as present in the virus), the polynucleotide sequences may further contain one or more alterations that lead to a codon that encodes the same amino acid as the codon given for the corresponding position in SEQ ID's 1, 3, 5 and/or 17, 19, 21, and this may even lead to a fully or totally artificial and/or synthetic sequence. Also, compared to nucleotide sequences given in SEQ ID's 1, 3, 5 and/or 17, 19, 21 and/or compared to any naturally occuring variant of these nucleic acid sequences (including DNA sequences corresponding to the RNA sequences as present in the virus), the polynucleotide sequences may further contain one or more alterations that lead to a conservative amino acid substitution, i.e. as mentioned above.

Providing a polynucleotide sequence that contains the desired alterations will be within the skill of the artisan and can involve techniques such as nucleic acid synthesis using an automated nucleic acid synthesis technique; introduction of (point)mutations into a nucleic acid that comprises the native 129 kD, 57 kD, and/or 186 kD sequences; and/or using or suitably combining parts or fragments of the 129 kD, 57 kD and/or 186 kD sequences, or any combination thereof. Also, in providing such a polynucleotide sequence, the skilled person may take into account the degeneracy of the genetic code and/or conservative amino acid substitutions, as mentioned above.

In order to provide a polynucleotide sequence that encodes a truncated replicase as defined above, a technique involving the introduction of a stopcodon into the native sequence is particularly preferred.

A particularly preferred technique of introducing the above alterations – including stopcodons - involves the use of a PCR reaction, in which the desired alterations are introduced into the amplified sequence(s) by the use of modified primers, i.e. primers that contain a suitable “mismatch” compared to the template sequence, leading to the desired alteration in the amplified sequence. This PCR-based technique may also be used to introduce one or more restriction sites into the amplified sequence in order to facilitate the cloning of the amplification products into the desired transformation vectors.

As further described in the Experimental Part, this may involve a single PCR-reaction, but may also involve two or more PCR reactions, each leading to a part of intended final sequence encoding the defective replicase in which the primers (e.g. with the desired alteration) form the ends of the fragments. These fragments may then be combined, for instance to provide a polynucleotide sequence that comprises a combination of such fragments, and/or to reconstitute the full 129 kD, 57 kD and/or 186 kD sequence, now containing the desired alteration compared to the native sequence, such as a stopcodon.

The PCR-reactions and the further steps following amplification, such as combining/joining the amplified sequences, can be carried out in a manner known per se, for instance as described in the Experimental Part and/or using the techniques described in US-A-4,683,202; Saiki et al., Science 239 (1988), 487-491 or PCR Protocols, 1990, Academic Press, San Diego, California, USA.

As the template for the PCR-reaction, a nucleic sequence encoding the native 129 kD, 57 kD and/or 186 kD sequence can be used, such as a cDNA derived from the native RNA sequence, or a plasmid containing such a sequence, including those described in the Experimental Part. The template used may itself already contain one or more alterations, compared to the corresponding native sequence.

As mentioned above, a preferred alteration involves the introduction of a stopcodon into the native 129 kD or 186 kD sequence, such that – upon transformation into a plant – the polynucleotide sequence thus obtained causes expression of a truncated replicase. In particular, such a stopcodon may be introduced into a sequence corresponding to the native 129 kD sequence, more in particular to that part of the native sequence that corresponds to the so-called GDD-motiv or to the so-called P-loop.

The polynucleotide sequence encoding the defective replicase is preferably in the form of - e.g. forms part of and/or is incorporated within - a genetic construct. The genetic construct is preferably a construct suitable for the transformation of a plant, plant cell and/or plant material, such as a plasmid, cosmid or vector, including co-integration vectors or binary vectors. The genetic construct may be DNA or RNA, and is preferably dsDNA.

Such a construct may further contain all known elements for genetic constructs, and in particular for genetic constructs intended for the transformation of plants, as long as the presence thereof does not interfere with the CGMMV resistance to be provided by the polynucleotide sequence encoding the defective replicase. Some non-limiting examples of such elements include leader sequences, terminators, enhancers, integration factors, selection markers, reporter genes, etc., and suitable elements will be clear to the skilled person.

These further elements may or may not be derived from plants, and may or may not be homologous to the plant that is to be transformed with the construct of the invention (hereinbelow referred to as the "*target plant*"). For instance, the further elements may also have been derived from micro-organisms, viruses, etc., and may also be elements that are natively associated with the CGMMV sequence. such as the native CGMMV leader sequence (5'-UTR sequence).

The nucleotide sequences encoding these further elements may have been isolated and/or derived from a naturally occuring source – for instance as cDNA – and/or from known available sources (such as available plasmids, etc.), and/or may have been provided synthetically using known DNA synthesis techniques.

For instance, a construct of the invention will usually contain a suitable promoter operatively linked to the polynucleotide sequence encoding the defective replicase, e.g. such that it is capable of directing the expression of the polynucleotide sequence. Suitable promoters can be chosen from all known constitutive, inducible, tissue specific or other promoters that can direct expression of a desired nucleotide sequence in a plant and/or in part of a plant, including specific tissues and/or individual cells of the plant. In particular, promoters are used that are suitable for use in species of the *Cucurbitaceae* family, such as cucumber.

A specifically preferred promoter is the plastocyanine-promoter. Use of the 35S promoter is less preferred, as it may be less reliable in cucumber.

The terminator can be any terminator that is effective in plants. A particularly preferred terminator is the nos-3' terminator.

The selection marker can be any gene that can be used to select – under suitable conditions such as the use of a suitable selection medium- plants, plant material and plant cells that contain –e.g. as the result of a succesful transformation - the genetic construct containing the marker. A particularly preferred selection marker is the *nptII*-gene, which can be selected for using kanamycin.

The construct of the invention further preferably contains a leader sequence. Any suitable leader sequence, including those of viral origin, can be used. Preferably, a leader sequence essentially identical to the 5'untranslated (5'-UTR) region of the CGMMV genome is used. This may be derived from the viral RNA, or may be provided synthetically, e.g. as described in the Experimental Part.

Although not preferred, the invention also encompasses constructs that encode a fusion of a defective replicase as mentioned above, and at least one further amino acid

sequence, such as a protein or polypeptide, or a part or fragment thereof. Preferably, expression of a defective replicase as (part of) such a fusion does not detract from the desired biological activity (i.e. protection against infection with CGMMV).

The construct of the invention can be provided in a manner known per se, which generally involves techniques such as restricting and linking nucleic acids/nucleic acid sequences, for which reference is made to the standard handbooks, such as Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd.ed.), Vols. 1-3, Cold Spring Harbor Laboratory (1989) of F. Ausubel et al, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987).

According to one embodiment, the genetic construct is preferably (also) in a form that can be stably maintained or inherited in a micro-organism, in particular a bacterium, more in particular a bacterium that can be used to transform a plant or plant material, such as *Agrobacterium*. In a further aspect, the invention also relates to such a micro-organism, in particular a bacterium, more in particular a bacterium that can be used to transform a plant, such as *Agrobacterium*, that contains a genetic construct according to the invention.

The genetic construct can be transformed into the target plant, plant cell or plant material by any suitable transformation technique known per se, including transformation with *Agrobacterium*, transformation with "denuded" DNA, for instance through particle bombardment or transformation of protoplasts through electroporation or treatment with PEG.

Examples of suitable vectors systems for use with *Agrobacterium* are for instance binary vectors such as pBI121 and derivatives thereof; co-integration vectors such as pGV1500 and derivatives of pBR322. Suitable systems for transformation with denuded DNA include *E.coli*-vectors with high copy number, such as pUC-vectors and pBluescript II (SK+) vectors.

Upon transformation, the construct may for instance be incorporated into the genomic DNA of the plant, or it may be maintained/inherited independantly in the plant (cell).

In a further aspect, the invention therefore comprises a method in which a plant, plant cell or plant material is transformed with a genetic construct as described above.

This method may also comprise cultivating the transformed plant cell or plant material into a mature plant, and may also comprise sexually or asexually reproducing or multiplying the transformed plant (and/or the mature plant obtained from the transformed plant cell or plant material).

The invention therefore also relates to a plant, plant cell or plant material, that has been transformed with – or more generally contains- a genetic construct as described above.

Preferably, such a plant, plant cell or such plant material is resistant against infection with CGMMV as described herein.

The invention furthermore relates to cultivation material such as seed, tubers, roots, stalks, seedlings etc. for such a plant, as well as descendants of such a plant, obtained through sexual or asexual reproduction techniques. Such cultivation material and/or descendants most preferably still contain or have inherited the genetic construct of the invention, and more preferably also are resistant against infection with CGMMV as described herein.

The invention will now be illustrated by means of the following non-limiting Experimental Part and by means of the Figures, in which:

- Figure 1 is a schematic representation of the genome of CGMMV;
- Figure 2 gives a phylogenetic tree of CGMMV-coat protein (cp) for CGMMV-SH and the ten European isolates, using the method of J. Hein with weighted residue table.
- Figures 3 - 8 show examples of some preferred genetic constucts of the invention, i.e those listed in Table 6 below.

Also, in the Experimental Part hereinbelow, enzymes, kits, etc. were usually used according to the instructions of the manufacturer and/or using well-established protocols, unless indicated otherwise.

Experimental Part

Example I: Cloning of the coat protein genes of 10 CGMMV-isolates

1. Collecting CGMMV isolates

To make use of coat protein-mediated protection (CPMP) strategy against CGMMV, it is necessary to clone the coat protein cistrons of the isolates, that are economically important. As the only sequence information available for CGMMV is derived from watermelon strains from the Far East, it was first decided to collect CGMMV isolates of important cucumber culture areas in Europe and the Mediterannean area. Table 1 lists the isolates collected from various geographical areas. All isolates were propagated on cucumber, and infected leaf material was stored at –80°C. The symptoms obtained after infection of cucumber cv. *Hokus* are listed in Table 1.

2. Design of PCR primers

The possibility of sequence divergence among the various collected isolates, and between the isolates and the published sequences of CGMMV-SH and CGMMV-W exists. In

order to identify nucleotide regions with a high degree of sequence conservation, that could serve as a basis of PCR primer design, an alignment study was carried out on corresponding sequences of CGMMV-SH, CGMMV-W and of some other related members of the tobamovirus group: Sunn-Hemp Mosaic Virus (SHMV, a variant of TMV) and Pepper Mild Mottle Virus (PMMV). For this purpose, a region of 800 nucleotides just 5' of the coat protein cistron and a region of 170 nucleotides forming the far 3' of the viral genome were compared. In this sequence alignment, regions with sufficient sequence homology among all compared viruses were identified. Based on these sequences, sets of PCR primers were designed, which are listed in Table 2.

Table 2. Design of primers for the RT-PCR amplification of coat protein sequences of CGMMV-isolates.

primers	sequence	position on CGMMV-SH sequence
5' primers		
97G01	AGGTGTCAGTGGAGAACTCATTGA	5004
97G02	GGCGTTGTGGTTTGTGG	5210
97G03	CTGTAGGGGTGGTGCTACTGT	5248
3' primer		
97G18	GCCCATAGAAACTTCAACGTC	6370

3. Amplification of the coat protein regions

From leaf material of cucumber plants infected with each of the 11 isolates described in Tabel 1, a total RNA extraction was prepared. Using each of the 5' primers listed in Table 2 in combination with 3' primer 97G18, reverse transcription of RNA and PCR amplification of cDNA with an annealing temperature of 55°C was established using a kit manufactured by Perkin Elmer Cetus. Especially in the reactions with the 5' primer 97G03 amplification products of the correct size were obtained for each of the 11 RNA samples. The PCR amplification products were directly cloned in T/A cloning vector pCR2.1 and introduced in *E.coli* strain INVαF'. For each of the RNA samples, the correct size of the cloned product (1.12 kb) was verified, and the clones were stored at –80°C. The amplification products of

CGMMV-isolates 1 to 10 cloned in pCR2.1 were designated pKG4301 to pKG4310, and the one of CGMMV-SH cloned in pCR2.1 was designated pKG4311.

4. Nucleotide sequence analysis of the coat protein cistrons

The sequences of the complete inserts of the plasmids pKG4301 to pKG4310 were determined by reading in both directions using m13 forward en m13 reverse sequencing primers. The sequence of the insert of pKG4311 was already known, as this plasmid contains a cDNA fragment of CGMMV-SH.

Sequence analysis confirmed, that in each case indeed the correct cDNA fragment of CGMMV had been obtained and cloned. With one exception, each amplified and cloned cDNA fragment consisted of 1123 base pairs, containing the CGMMV coat protein cistron and a large part of the CGMMV movement protein cistron.

The cloned sequences of all collected European isolates (isolates 1 to 10) are approximately 97% homologous among each other, but differ on average by 10% from the published sequence of CGMMV-SH. Comparison of each individual sequence revealed, that isolates 1 and 2, both from Eastern Europe, are extremely alike. The same very high degree of identity was found between both isolates from cucumber greenhouses in the Netherlands (isolates 4 and 5) and between both isolates obtained from the Almería area in Spain (isolates 9 and 10). None of the cDNA sequences was 100% identical to any of the other ones, but the differences in sequence are no more than a few nucleotides, and sometimes only one nucleotide in the coding region of the coat protein cistron. The Japanse isolate CGMMV-SH is clearly different from any of the European isolates.

5. Coat protein amino acid sequence analysis

Based on the nucleotide sequences of the Open Reading Frames (ORF) of the coat protein cistrons of the 10 isolates, the amino acid sequence could be deduced. In each of the analyzed sequences, the ORF consisted of a region of 486 nucleotides, coding for a protein of 161 amino acid residues. The predicted molecular mass of this protein is 17.3 kD, corresponding to earlier published results. The homology among the predicted protein sequences of the various isolates is as high as 98.1%. The only deviations are found for amino acid residue 19 (usually valine), residue 65 (mostly serine) and residue 84 (mostly leucine).

The sequence of the coat protein of the Japanse isolate CGMMV-SH only differs by 1 amino acid (residue 65) from the consensus sequence.

Example II: Cloning of the replicase gene of CGMMV

1. Strategy for replicase-mediated protection

By way of example, two approaches to replicase-mediated protection (RMP) against virus infections in plants were investigated.

One approach makes use of defective replicase genes in the form of truncated Open Reading Frames (ORF), in which the sequence downstream from the GDD motif had been truncated or altered through mutation.

The other approach makes use of the expression of the 'read-through' part of the replicase gene, i.e. the 57 kD sequence. It is thought that this ORF is not translated in the plant cell, but forms part of a larger 'read-through' ORF combining the coding regions of both the 129 kD replicase gene and the putative 57 kD protein gene, resulting in a protein of 189 kD. However, expressing merely the 57 kD protein ORF in plant cells may result in a extremely strong resistance to infection by both virus particles and viral RNA, which also would be capable of resisting high temperatures, as well as high inoculum concentrations.

For either of these approaches, either the full-length CGMMV replicase gene must be cloned, or both constituting parts must be cloned separately.

2. Design of primers

Because of the high sequence homology of the coat protein genes of 11 CGMMV isolates it was assumed, that the sequences of the replicase genes of the various isolates would be also highly conserved. Based on the complete sequence of the CGMMV-SH isolate, primers were designed for the PCR amplification of the 57 kD ORF and of the 129 kD ORF (Tables 3 and 4). The primers were designed such, that they contain restriction sites for the future cloning of the amplification products. The 5' primers contain an *Nco*I-site positioned such, that it will coincide with the ATG start codon of the amplified ORF. The 3' primers contain a *Sac*I-site downstream from the stop codon.

Table 3. Design of primers for the LR-RT-PCR amplification of the 57 kD replicase sequence of CGMMV.

primers	sequence	position on CGMMV-SH sequence
5' primer		
98A88	CCATGGAGAATTCGCTGTATGTC C	3497
3' primer		
98A86	CGAGCTCTCGACTGACACCTTAC	5001

Table 4. Design of primers for the LR-RT-PCR amplification of the 129 kD replicase gene sequence of CGMMV.

primers	sequence	position on CGMMV-SH sequence
5' primers		
98A84	CCATGGCAAACATTAATGAAC	59
98A85	CAACCATGGCAAACATTAATG	56
3' primer		
98G63	TAACAGGGAGGAAAATATTACG	

3. Long-Range Reverse Transcriptase Polymerase Chain Reactions

From the known sequence of CGMMV-SH it was derived that the size of the 57 kD protein gene is 1.5 kb. Such a size is at the limit of the size range that can be amplified in a PCR with standard *Taq* polymerase. For the amplification of this cDNA fragment, and certainly for the amplification of the cDNA fragment for the 129 kD replicase gene, a different polymerase suitable for long range amplifications must be used. In these experiments, rTth DNA polymerase was used.

For direct amplification of cDNA fragments from total RNA extractions a RT-PCR kit is normally employed, combining in one reaction the activity of the Reverse Transcriptase (RT), producing a single cDNA strand complementary to the RNA template strand beginning

at one primer annealed to the 3' end of the RNA molecules, and the activity of the Polymerase, amplifying the thus produced single stranded cDNA molecule in a normal PCR fashion.

Because of the need to use long range polymerase. it was attempted to combine the RT with the long range polymerase to produce in one reaction large-size amplification products directly from total RNA extracts. This type of reaction was called a Long Range Reverse Transcriptase Polymerase Chain Reaction (LR-RT-PCR).

4. LR-RT-PCR amplification and cloning of the 57 kD protein gene

Using the primers listed in Table 3 and the LR-RT-PCR described above, a specific 1.5 kb amplification product was obtained from total RNA extracts of cucumber leaves infected with CGMMV-4. This isolate was chosen, as it originated from the Dutch cucumber greenhouse cultures, and would thus represent an economically important isolate. Because long range polymerases contain a 'proof reading' activity and do not leave A-additions on the amplification products, as does the *Taq* polymerase normally employed in PCR, direct cloning of the amplification products in a TA vector accomodating the A-additions was not possible. Therefore, the amplification products were briefly treated with *Taq* polymerase, resulting in the addition of A-overhangs on the amplified DNA molecules. These molecules could then easily be cloned in the TA vector pCR2.1, and transformed to *E.coli* MC1061. Clones with the correct insert size of 1.5 kb were stored at –80°C and are known as pKG4321.

5. Sequence analysis of the 57 kD protein gene

The nucleotide sequence of the cloned insert of pKG4321 was determined by double-stranded sequencing using m13 forward en m13 reverse primers and subsequent primer walking steps. The ORF coding for a putative 57 kD protein gene (SEQ ID no 3) showed 90% homology at the nucleotide level to the corresponding sequence of the Japanse isolate CGMMV-SH (SEQ ID no 19). The predicted amino acid sequence (SEQ ID no 4) shows a 98.2% homology to the one predicted by the CGMMV-SH sequence (SEQ ID no 20). The GDD motif characteristic for viral replicase genes resides at amino acid residues 364-366.

6. LR-RT-PCR amplification and cloning of the 129 kD replicase gene

Using the primers listed in Table 4 in a Long Range Reverse Transcriptase Polymerase Chain Reaction as described under 3 , one specific amplification product of 3.5 kb

representing the viral 129 kD replicase gene was obtained from total RNA of cucumber leaves infected with CGMMV isolate 4 (Table 1). Because long range polymerases contain a 'proof reading' activity and do not leave A-additions on the amplification products, as does the *Taq* polymerase normally employed in PCR, direct cloning of the amplification products in a TA vector accomodating the A-additions was not possible. Therefore, the amplification products were briefly treated with *Taq* polymerase, resulting in the addition of A-overhangs on the amplified DNA molecules. These molecules could easily be cloned in the TA vector pCR2.1, and transformed to *E.coli* MC1061. Clones with the correct insert size of 3.5 kb were stored at –80°C and are known as pKG4322.

7. Sequence analysis of the 129 kD protein gene

The nucleotide sequence of the amplification product cloned in pKG4322 was determined by double-stranded sequencing using m13 forward en m13 reverse primers, and a primer walking strategy. The ORF coding for the 129 kD replicase gene (SEQ ID no 1) showed 88% homology at the nucleotide level to the corresponding sequence of the Japanse isolate CGMMV-SH (SEQ ID no 17). The ORF of the Dutch cucumber greenhouse isolate codes for a replicase protein of 1144 amino acids, which is one amino acid in extra in comparison to the CGMMV-SH strain. The predicted amino acid sequence (SEQ ID no. 2) shows a 97.1% homology to the one predicted by the CGMMV-SH sequence (SEQ ID no 18).

Two GDD motifs are found at amino acid residues 256-258 and 540-542.

8. Site-directed mutagenesis of the 129 kD ORF

As explained above, one approach to obtain RMP in plant cells was to make use of replicase genes truncated either in the GDD motif, or truncated in the P-loop of the helicase domain. In order to create gene expression cassettes carrying such truncated genes, a site-directed mutagenesis approach was followed to introduce stop codons at the required positions in the ORF. To this end, several parts of the 129 kD replicase ORF were re-amplified from pKG4322 as a template using specifically designed primers that included unique restriction sites for future re-assembling of the thus amplified products, as well as the required mutations in the form of stop codons (Table 5). These stop codons should ensure the proper truncation of the translation of the protein. Several stop codons were designed one after the other in the three reading frames in these primers, thus ensuring an effective translation-deficient mutation.

Table 5. Design of primers for the site-directed mutagenesis of the 129 kD replicase gene of CGMMV.

primers	sequence	
98L99	GAGCTCGGATCCACTAGTAACGGC	
98L107	TAGAGCTCTTGAAGCTAAGCAAATTCCG	
98L108	TTCAAGAGCTCTAATCACCGAAGACAAAGGC	
98L102	GAATTATATCGATTATCTATCGGC	
98L103	GATAATCGATATAATTCTTCATCTGCC	
98L104	AACTAGTAATTGATGATCTGTTCAAGAAG	
98L105	AATTACTAGTTTCCGGAAGCAAGCAGCTCAG	
98L106	GCCCTCTAGATGCATGCTCGAG	

Using primers 98L103 and 98L104, a fragment from the downstream half of the 129 kD gene from the GDD motif up to the *Cla*I-site was amplified, while simultaneously stop codons were introduced at the site of the GDD motif. This fragment cloned in TA vector pCR2.1 was called pKG4325.

Using primers 98L105 and 98L106, a fragment corresponding to the 5' half of the 129 kD gene up to the GDD motif was amplified, while simultaneously a stop codon was introduced at the site of the GDD motif. This fragment cloned in T/A vector pCR2.1 was called pKG4326.

Replacing an *Xba*I-*Cla*I fragment of pKG4322 with the combined amplified products of pKG4325 and pKG4326 reconstitutes the full-length 129 kD replicase ORF of pKG4322 with stop codons introduced at the site of the GDD motif. This construct is named pKG 4329.

Using primers 98L99 and 98L107, a fragment at the far downstream end of the 129 kD gene from the P-loop to the end of the ORF was amplified, while simultaneously stop codons were introduced at the site of the P-loop. This fragment cloned in T/A vector pCR2.1 was called pKG4327.

Using primers 98L108 and 98L102, a fragment corresponding to a central part of the 129 kD gene from the GDD motif up to the P-loop was amplified, while simultaneously a stop codon was introduced at the site of the P-loop. This fragment cloned in T/A vector pCR2.1 was called pKG4328.

Replacing an *Bam*HI-*Cla*I fragment of pKG4322 with the combined amplified products of pKG4327 and pKG4328 reconstitutes the full-length 129 kD replicase ORF of pKG4322 with stop codons introduced at the site of the P-loop. This construct is named pKG4330.

Example III: Transformation of cucumber

1. Construction of a CGMMV-leader sequence

For optimal expression and stability of the replicase gene transcripts in plant cells, it was thought necessary to add a sequence identical to the 5' untranslated (5' UTR) region of the CGMMV genome upstream from the ORF sequence in the plant expression cassette. Because the 5' UTR of viral genomes contain highly repetitive RNA, this sequence could not be obtained by RT-PCR amplification, as no specific primers could be designed. Instead, a synthetic region identical to the 5' UTR of CGMMV-SH was assembled from the four oligonucleotide sequences:
97G40 (CTAGAGTTTTAATTTTTATAATTAAACAAA),
97G41 (TCAAAATTAAAAATATTAATTTGTTTGTTGTTGTTG),
97G42 (CAACAACAACAACAACAAACAATTTTAAAACAACAC) and
97G43 (TTGTTGTTTGTTAAAATTTTGTTGTGGTAC).

These oligonucleotides were designed such, that outside the sequence corresponding to the 5' UTR they contain restriction sites for *Xba*I and *Nco*I, thus facilitating further cloning. Adding the four oligonucleotides together will cause spontaneous assembling due to the design of extensive regions of overhang. Using these restriction sites, the assembled mixture was cloned in a plant expression vector containing an *Arabidopsis thaliana* plastocyanin promoter (Vorst *et al.*, 1993) and a *Agrobacterium tumefaciens* nopaline synthase terminator sequence (Depicker *et al.*, 1982) in a pUC19-derived plasmid. This expression cassette was called pKG1315. The complete expression cassette consisting of the plastocyanin promoter, the CGMMV leader sequence and the *nos* terminator was subsequently removed from pKG1315 using *Hin*dIII and *Eco*RI as restriction enzymes, and recloned in the corresponding restriction sites of:

1) an intermediate type *Agrobacterium* transformation vector for coïntegrate type vector systems containing an *npt*II selectable marker gene cassette to create pKG1575, and
2) an intermediate type *Agrobacterium* transformation vector for coïntegrate type vector systems containing no selectable marker gene to create pKG1110.

2. Construction of transformation vectors

The three cloned and modified replicase constructs of pKG4321, pKG4329 and pKG4330 were isolated from the plasmids by restriction with *Bam*HI (filled in with Klenow) and *Nco*I and ligated into the *Sac*I (filled in with Klenow) and *Nco*I sites of each of the two transformation vectors pKG1575 and pKG1110, resulting in a total of six transformation vectors, listed in Table 6.

Table 6. List of six transformation vectors for the expression in plants of parts of the CGMMV-replicase gene.

vector	vector type	modified CGMMV-replicase gene
pKG4331	intermediate type with nptII	57 kD ORF
pKG4332	intermediate type with nptII	129 kD ORF with stopcodon in GDD motif
pKG4333	intermediate type with nptII	129 kD ORF with stopcodon in P-loop
pKG4334	intermediate type	57 kD ORF
pKG4335	intermediate type	129 kD ORF with stopcodon in GDD motif
pKG4336	intermediate type	129 kD ORF with stopcodon in P-loop

3. Transformation of cucumber

The intermediate type transformation vectors pKG4331 and pKG4333 were introduced into *Agrobacterium tumefaciens* strain GV2260 by tri-parental mating. Transconjugants which had incorporated the intermediate type vector into their Ti-plasmids through homologous recombination were selected on the basis of streptomycin and spectinomycin resistance and analyzed for the correct insertion of the vector.
Cucumber plants were transformed with these two strains of *Agrobacterium*, as well as with an *Agrobacterium* strain harbouring only the *npt*II selection marker, using published procedures. A number of transgenic cucumber plants were obtained. The plants were transferred to a greenhouse to flower and set seed. The seedlings germinating from these R1 seed were mechanically infected with CGMMV isolate 1-3 weeks post-inoculation, the plants were scored for symptoms of virus infection, as described in the assay for tolerance to virus infection set out in under 4. below.

4. Assay for tolerance to virus infection

The seedlings of transgenic cucumber germinating from these R1 seed were mechanically infected with CGMMV isolate 1. Fresh inoculum was prepared from a crude leaf extract of susceptible non-transgenic cucumber plants cv. Hokus pre-infected with this same isolate 3 weeks previously. Seedlings of non-transformed cucumber plants were used as controls in the assay. During 21 days post-inoculation the appearance of viral symptoms was scored visually every 2 days. In this assay, individual plants are scored as being tolerant when they remain free of visible symptoms for at least 7 days, and preferably more than 14 days, and more preferably more than 21 days post-inoculation.

Sixty-four independent transgenic lines were assayed, with 14 to 20 seedlings for each line. Control seedlings all became diseased within 9 days post-inoculation. A number of seedlings in seventeen of the transgenic lines showed clear absence of symptoms for a prolonged period of time, and remained free of symptoms after 21 days post-inoculation. Of some transgenic lines, the number of symptom-free plants coresponded to Mendelian segregation of a transgene present in a single locus. In one particular transgenic cucumber line 4 out of 14 seedlings remained symptom-free during the assay period, which may indicate that the tolerant phenotype corresponds to the homozygous state of a transgene present in one single locus, although, as mentioned above, the invention is not limited to a specific mechanism.

Hereinabove, the invention has been described under the assumption that resistance against CGMMV is generated "at the protein level", i.e. that the "*nucleotide sequence encoding a defective variant of the replicase gene of CGMMV*" codes for a "*defective replicase*", the expression of which at cellular level generates the desired resistance against CGMMV. However, the invention is not limited to any explanation or mechanism, and is not particularly limited to the use of a particular type of nucleotide sequence (i.e. encoding a "defective replicase"). For instance, an alternative mechanism is based on the assumption that the resistance to CGMMV is generated at the RNA-level, e.g. downregulation of gene expression due to RNA sequence homology. The use of any nucleotide sequence that can be used to provide resistance against CGMMV and that at the same time does not provide replicase activity, is envisaged by the invention.

References

- Anderson, J.M., Palukaitis, P. and M. Zaitlin (1992) A defective replicase gene induces resistance to cucumber mosaic virus in transgenic tobacco. Proc.Nat.Acad.Sci.USA 89: 8759-8763.
- Audy *et al.* (1993) Molec.Plant-Microbe Interact. 7: 15-23.
- Beachy, R.N., S. Loesch-Fries and N.E. Tumer (1990) Coat-protein mediated resistance against virus infection. Ann.Rev.Phytopathology 28: 451-474.
- Christou *et al.* (1992) IAPTC Newsletter 2-14.
- Cooper *et al.* (1995) Virology 206: 307-313.
- Depicker *et al.* (1982) J.Mol.Appl.Genet. 561-573.
- Donson, J., C.M. Kearney, T.H. Turpen, I.A.Khan, G. Kurath, A.M. Turpen, G.E. Jones, W.O. Dawson and D.J. Lewandowski (1993) Broad Resistance to Tobamoviruses is mediated by a modified Tobacco Mosaic Virus Replicase Transgene. Mol.Plant-Microbe Interact. 6: 635-642.
- Goldbach, R. and P. De Haan (1993) Prospects of engineered forms of resistance against tomato spotted wilt virus. Seminars in Virology 4: 381-387.
- Golemboski, D.B., G.P. Lomonosoff and M. Zaitlin (1990) Plants transformed with a toacco mosaic virus nonstructural gene sequence are resistant to the virus. Proc. Natl. Acad. Sci.USA 87: 6311-6315.
- Hollings, M., Y. Komuro and H. Tochihara (1975) Cucumber Green Mottle Mosaic Virus. C.M.I./A.A.B. Descriptions of Plant Viruses, nr. 154.
- Horsch, R.B., J.G. Fry, N.L. Hoffmann, D. Eichholtz, S.G. Rogers and R.T. Fraley (1985) A simple and general method for transferring genes into plants. Science 227: 1229-1231.
- Kunik, T., R. Salomon, D. Zamir, N. Navot, M. Zeidan, I. Michelson, Y. Gafni and H. Czosnek (1994) Transgenic tomato plants expressing the tomato yellow leaf curl virus capsid protein are resistant to the virus. Bio/Technology 12: 500-504.
- Meshi, T., R. Kiyama, T. Ohno and Y. Okada (1983) Nucleotide sequence of the coat protein cistron and the 3' noncoding region of cucumber green mottle mosaic virus (watermelon strain) RNA. Virology 127: 54-64.
- Prins *et al.* (1994) Molec.Plant-Microbe Interact. 8: 85-91.
- Ugaki, M., M. Tomiyama, T. Kakutani, S. Hidaka, T. Kiguchi, R. Nagata, T. Sato, F. Motoyoshi and M. Nishiguchi (1991) The complete nucleotide sequence of Cucumber Green Mottle Mosaic Virus (SH strain) genomic RNA. J.Gen.Virol. 72: 1487-1495.

- Rogers, S.G., H.J. Klee, R.B. Horsch and R.T. Fraley (1986) Gene transfer in plants: production of transformed plants using Ti plasmid vectors. Meth. Enzymol. 118: 627-640.
- Shillito, R.D., M.W. Saul, J. Paszkowski, M. Muller and I. Potrykus (1985) High efficiency direct gene transfer to plants. Bio/Technology 3: 1099-1103.
- Vaira, A.M., L. Semeria, S. Crespi, V. Lisa, A. Allavena and G.P. Accotto (1995) Resistance to Tosposviruses in Nicotiana benthamiana transformed with the N gene of tomato spotted wilt virus: correlation between transgene expression and protection in primary transformants. Molec.Plant-Microbe Interact. 8: 66-73.
- Vorst, O., P. Kock, A. lever, B. Weterings, P. Weisbeek and S. Smeekens (1993) The promoter of the Arabidopsis thaliana plastocyanin gene contains a far upstream enhancer-like element involved in chloroplast-dependent expression. Plant Journal 4: 933-945.
- Wilson, T.M.A. (1993) Strategies to protect crop plants against viruses: pathogen-derived resistance blossoms. Proc.Natl.Acad.Sci.USA 90: 3143-3141.

CLAIMS

1. Method for generating resistance in a plant or in a plant cell against infection with CGMMV, said method comprising at least the step of transforming said plant or plant cell with a polynucleotide sequence that

i) upon (at least) transformation into a plant and transcription into RNA generates resistance against infection with CGMMV in said plant; and

ii) upon (at least) transformation into a plant and transcription into RNA does not lead to generation of (any) replicase activity in said plant.

2. Method for providing a transgenic plant and/or a transgenic plant cell that is resistant against infection with CGMMV, comprising at least the step of transforming said plant or plant cell with a polynucleotide sequence that

i) upon (at least) transformation into a plant and transcription into RNA generates resistance against infection with CGMMV in said plant; and

ii) upon (at least) transformation into a plant and transcription into RNA does not lead to generation of (any) replicase activity in said plant.

3. Method according to claim 1 and/or 2, in which the polynucleotide sequence is a nucleotide sequence that comprises, and preferably consists of:

- a nucleotide sequence corresponding to the sequence given in SEQ ID no.1, to the sequence given in SEQ ID no.17 or to the nucleotide sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.1, SEQ ID no. 17 and/or the naturally occuring variant thereof - one or more nucleotides have been added, replaced and/or removed;
- a nucleotide sequence corresponding to the sequence given in SEQ ID no.5, to the sequence given in SEQ ID no.21, to the nucleotide sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.5, SEQ ID no. 21 and/or the naturally occuring variant thereof - one or more nucleotides have been added, replaced and/or removed;
- a nucleotide sequence corresponding to the sequence given in SEQ ID no.3, to the sequence given in SEQ ID no.19 or to the nucleotide sequence of a naturally occuring variant thereof;

– a nucleotide sequence corresponding to the sequence given in SEQ ID no.3, to the sequence given in SEQ ID no.19 or to the nucleotide sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.3, SEQ ID no. 19 and/or the naturally occuring variant thereof - one or more nucleotides have been added, replaced and/or removed;

such that said nucleotide sequence is capable, upon (at least) transformation into a plant and transcription into RNA, to confer to said plant resistance against infection with CGMMV, and such that said nucleotide sequence, upon (at least) transformation into a plant and transcription into RNA, is not capable of generating of (any) replicase activity in said plant.

4. Method according to any of claims 1-3, in which the polynucleotide sequence encodes a polypeptide or protein that

i) upon being expressed in a plant is capable of generating resistance against CGMMV in said plant; and

ii) upon being expressed in a plant has no replicase activity.

5. Method according to claim 4, in which the polynucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of:

– an amino acid sequence corresponding to the sequence given in SEQ ID no.2, to the sequence given in SEQ ID no.18 or to the amino acid sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.2, SEQ ID no.18 and/or the naturally occuring variant thereof - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed;

– an amino acid sequence corresponding to the sequence given in SEQ ID no.6, to the sequence given in SEQ ID no.22 or to the amino acid sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.6, SEQ ID no. 22 and/or the naturally occuring variant thereof - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed

– an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no.20 or to the amino acid sequence of a naturally occuring variant thereof;

- an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no.20 or to the amino acid sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.4, SEQ ID no. 20 and/or the naturally occuring variant thereof - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed;

or any combination thereof, provided that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

6. Method according to claim 4 or 5, in which the polynucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of :

- an amino acid sequence corresponding to a part or fragment of the sequence given in SEQ ID no.2, to a part or fragment of the sequence given in SEQ ID no.18 and/or to a part of fragment of the amino acid sequence of a naturally occuring variant thereof, or corresponding to a combination of two or more such parts or fragments;
- an amino acid sequence corresponding to a part or fragment of the sequence given in SEQ ID no.6, to a part or fragment of the sequence given in SEQ ID no.22 and/or to a part of fragment of the amino acid sequence of a naturally occuring variant thereof, or corresponding to a combination of two or more such parts or fragments; ts;
- an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no.20, and/or to the amino acid sequence of a naturally occuring variant thereof;

such that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

7. Method according to claim 4 or 5, in which the polynucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of :

- an amino acid sequence corresponding to a truncated variant of the sequence given in SEQ ID no.2, to a truncated variant of the sequence given in SEQ ID no.18 and/or to a truncated variant of the amino acid sequence of a naturally occuring variant thereof;

- an amino acid sequence corresponding to a truncated variant of the sequence given in SEQ ID no.6, to a truncated variant of the sequence given in SEQ ID no.22 and/or to a truncated variant of the amino acid sequence of a naturally occuring variant thereof;
- an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no 20, and/or to the amino acid sequence of a naturally occuring variant thereof;

such that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

8. Method according to claim 7, in which the polynucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of, the amino acid sequence given in SEQ ID no.2, the amino acid sequence given in SEQ ID no 18 and/or the amino acid sequence of a naturally occuring variant thereof, that has been truncated in the GDD-motif or in the P-loop.

9. Method according to any of the preceding claims, further comprising at least one step of cultivating the transformed plant cell into a mature plant

10. Method according to any of the preceding claims, further comprising at least one step of sexually or asexually reproducing or multiplying the transformed plant and/or the mature plant obtained from the transformed plant cell of claim 9.

11. Method according to any of the preceding claims, in which the plant is a plant that is susceptible to infection with CGMMV, more preferably a plant belonging to the *Cucurbitaceae* family, such as melon (*Cucumis melo*), cucumber (*C.sativus*), watermelon (*Citrullus vulgaris*) and bottlegourd (*Lagenaria siceraria*)

12. Genetic construct suitable for transforming a plant, said construct at least comprising nucleotide sequence that

i) upon (at least) transformation into a plant and transcription into RNA generates resistance against infection with CGMMV in said plant; and

ii) upon (at least) transformation into a plant and transcription into RNA does not lead to generation of (any) replicase activity in said plant;

and optionally comprising further elements of genetic constructs known per se.

13. Genetic construct according to claim 12. at least comprising a nucleotide sequence that comprises, and preferably consists of:

- a nucleotide sequence corresponding to the sequence given in SEQ ID no.1, to the sequence given in SEQ ID no.17 or to the nucleotide sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.1, SEQ ID no. 17 and/or the naturally occuring variant thereof - one or more nucleotides have been added, replaced and/or removed;
- a nucleotide sequence corresponding to the sequence given in SEQ ID no.5, to the sequence given in SEQ ID no.21, to the nucleotide sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.5, SEQ ID no. 21 and/or the naturally occuring variant thereof - one or more nucleotides have been added, replaced and/or removed;
- a nucleotide sequence corresponding to the sequence given in SEQ ID no.3, to the sequence given in SEQ ID no.19 or to the nucleotide sequence of a naturally occuring variant thereof;
- a nucleotide sequence corresponding to the sequence given in SEQ ID no.3, to the sequence given in SEQ ID no.19 or to the nucleotide sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.3, SEQ ID no. 19 and/or the naturally occuring variant thereof - one or more nucleotides have been added, replaced and/or removed;

such that said construct is capable, upon (at least) transformation into a plant and transcription into RNA, to confer to said plant resistance against infection with CGMMV, and such that said construct, upon (at least) transformation into a plant and transcription into RNA, is not capable of generating of (any) replicase activity in said plant.

14. Genetic construct according to claim 12 or 13, at least comprising a nucleotide sequence that encodes a polypeptide or protein that

i) upon being expressed in a plant is capable of generating resistance against CGMMV in said plant; and

ii) upon being expressed in a plant has no replicase activity.

15. Genetic construct according to claim 14, in which the nucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of:

- an amino acid sequence corresponding to the sequence given in SEQ ID no.2, to the sequence given in SEQ ID no.18 or to the amino acid sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.2, SEQ ID no.18 and/or the naturally occuring variant thereof - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed;
- an amino acid sequence corresponding to the sequence given in SEQ ID no.6, to the sequence given in SEQ ID no.22 or to the amino acid sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.6, SEQ ID no. 22 and/or the naturally occuring variant thereof - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed
- an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no.20 or to the amino acid sequence of a naturally occuring variant thereof;
- an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no.20 or to the amino acid sequence of a naturally occuring variant thereof, in which - compared to the sequence of SEQ ID no.4, SEQ ID no. 20 and/or the naturally occuring variant thereof - one or more amino acids have been added, replaced or removed, preferably replaced or removed, more preferably removed;

or any combination thereof, provided that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

16. Genetic construct according to claim 14 or 15, in which the nucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of :

- an amino acid sequence corresponding to a part or fragment of the sequence given in SEQ ID no.2, to a part or fragment of the sequence given in SEQ ID no.18 and/or to a part of fragment of the amino acid sequence of a naturally occuring variant thereof, or corresponding to a combination of two or more such parts or fragments;

- an amino acid sequence corresponding to a part or fragment of the sequence given in SEQ ID no.6, to a part or fragment of the sequence given in SEQ ID no.22 and/or to a part of fragment of the amino acid sequence of a naturally occuring variant thereof, or corresponding to a combination of two or more such parts or fragments; ts;
- an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no.20, and/or to the amino acid sequence of a naturally occuring variant thereof;

such that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

17. Genetic construct according to any of claims 14-16, in which the nucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of :

- an amino acid sequence corresponding to a truncated variant of the sequence given in SEQ ID no.2, to a truncated variant of the sequence given in SEQ ID no.18 and/or to a truncated variant of the amino acid sequence of a naturally occuring variant thereof;
- an amino acid sequence corresponding to a truncated variant of the sequence given in SEQ ID no.6, to a truncated variant of the sequence given in SEQ ID no.22 and/or to a truncated variant of the amino acid sequence of a naturally occuring variant thereof;
- an amino acid sequence corresponding to the sequence given in SEQ ID no.4, to the sequence given in SEQ ID no 20, and/or to the amino acid sequence of a naturally occuring variant thereof;

such that the resulting protein or polypeptide shows no replicase activity, but is still capable - upon expression in a plant - to generate resistance against CGMMV in said plant.

18. Genetic construct according to claim 17, in which the nucleotide sequence encodes a protein or polypeptide that comprises, and preferably consists of, the amino acid sequence given in SEQ ID no.2, the amino acid sequence given in SEQ ID no.18 and/or the amino acid sequence of a naturally occuring variant thereof, that has been truncated in the GDD-motif or in the P-loop.

19. Genetic construct according to any of claims 12-18, in which the nucleotide sequence is under control of the plastocyanine-promoter.

20. Genetic construct according to any of claims 12-19, in which the nucleotide sequence is preceded by the native CGMMV leader (5'-UTR) sequence.

21. Genetic construct according to any of claims 12-20, in a form that can be stably maintained or inherited in a micro-organism, in particular a bacterium, more in particular a bacterium that can be used to transform a plant or plant material, such as *Agrobacterium*.

22. Micro-organism, in particular bacterium, more in particular a bacterium that can be used to transform a plant, such as *Agrobacterium*, that contains a genetic construct according to any of claims 12-21, and in particular according to claim 21.

23. Transgenic plant or plant cell, obtainable or obtained by a method according to one of claims 2-11, or a descendant of such a plant.

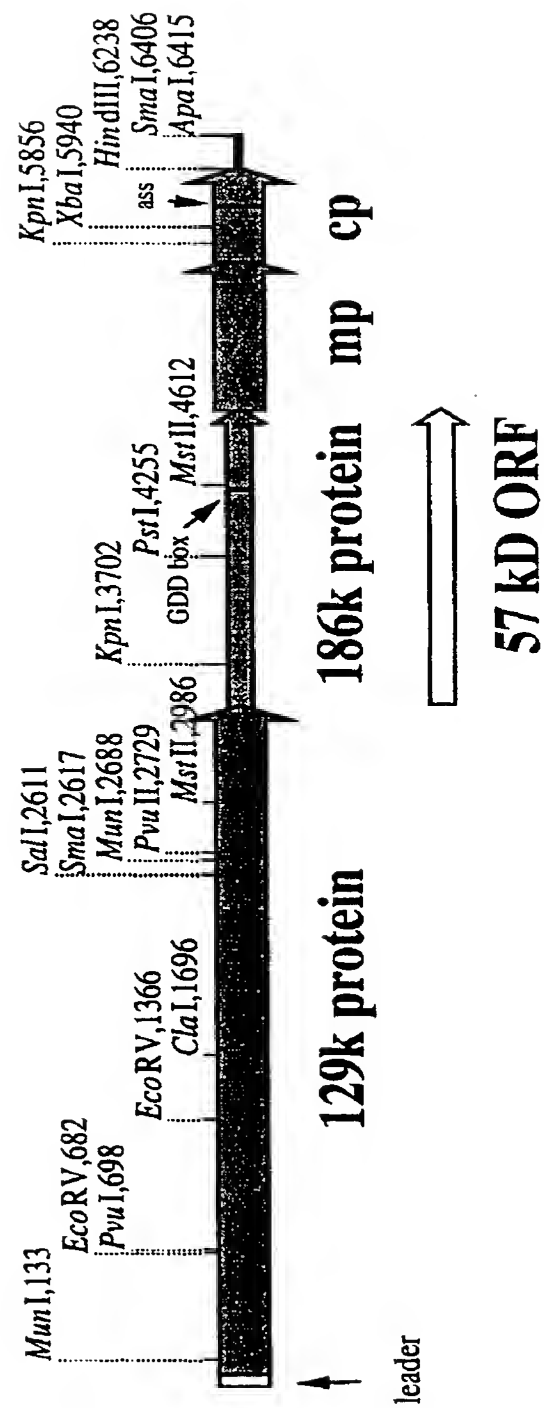
24. Plant, plant cell or plant material that has been transformed with genetic construct according to any of claims 12-21, or a descendant of such a plant.

25. Plant according to claim 24 or 25, being a plant that is susceptible to infection with CGMMV, more preferably a plant belonging to the *Cucurbitaceae* family, such as melon (*Cucumis melo*), cucumber (*C.sativus*), watermelon (*Citrullus vulgaris*) and bottlegourd (*Lagenaria siceraria*)

26. Cultivation material such as seed, tubers, roots, stalks, seedlings for a plant according to claim 23, 24 or 25.

Fig 1

6421 nucleotides



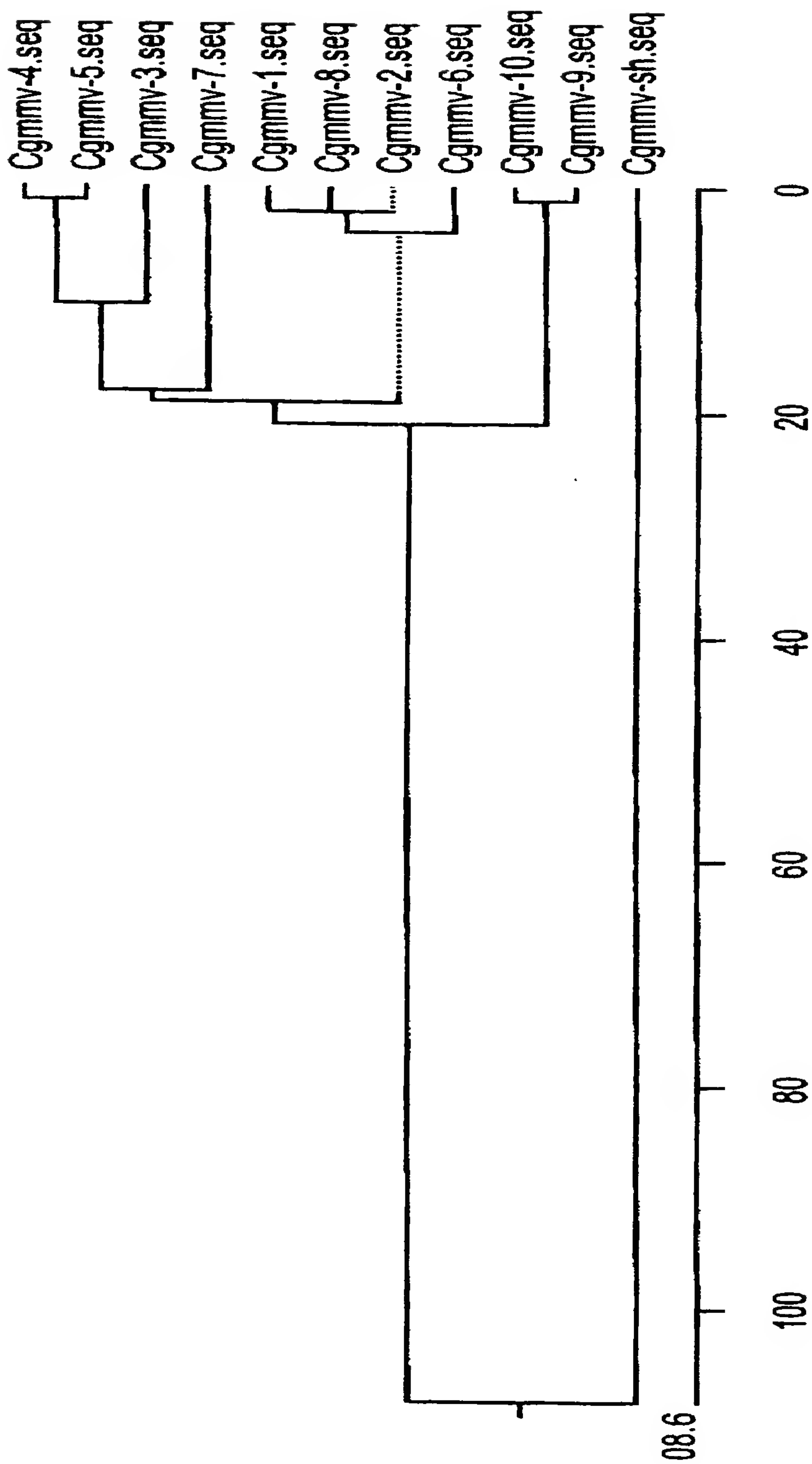

Cgmmv-4.seq
Cgmmv-5.seq
Cgmmv-3.seq
Cgmmv-7.seq
Cgmmv-1.seq
Cgmmv-8.seq
Cgmmv-2.seq
Cgmmv-6.seq
Cgmmv-10.seq
Cgmmv-9.seq
Cgmmv-sh.seq
0
20
40
60
80
100
108.6

Fig 2

Fig 3

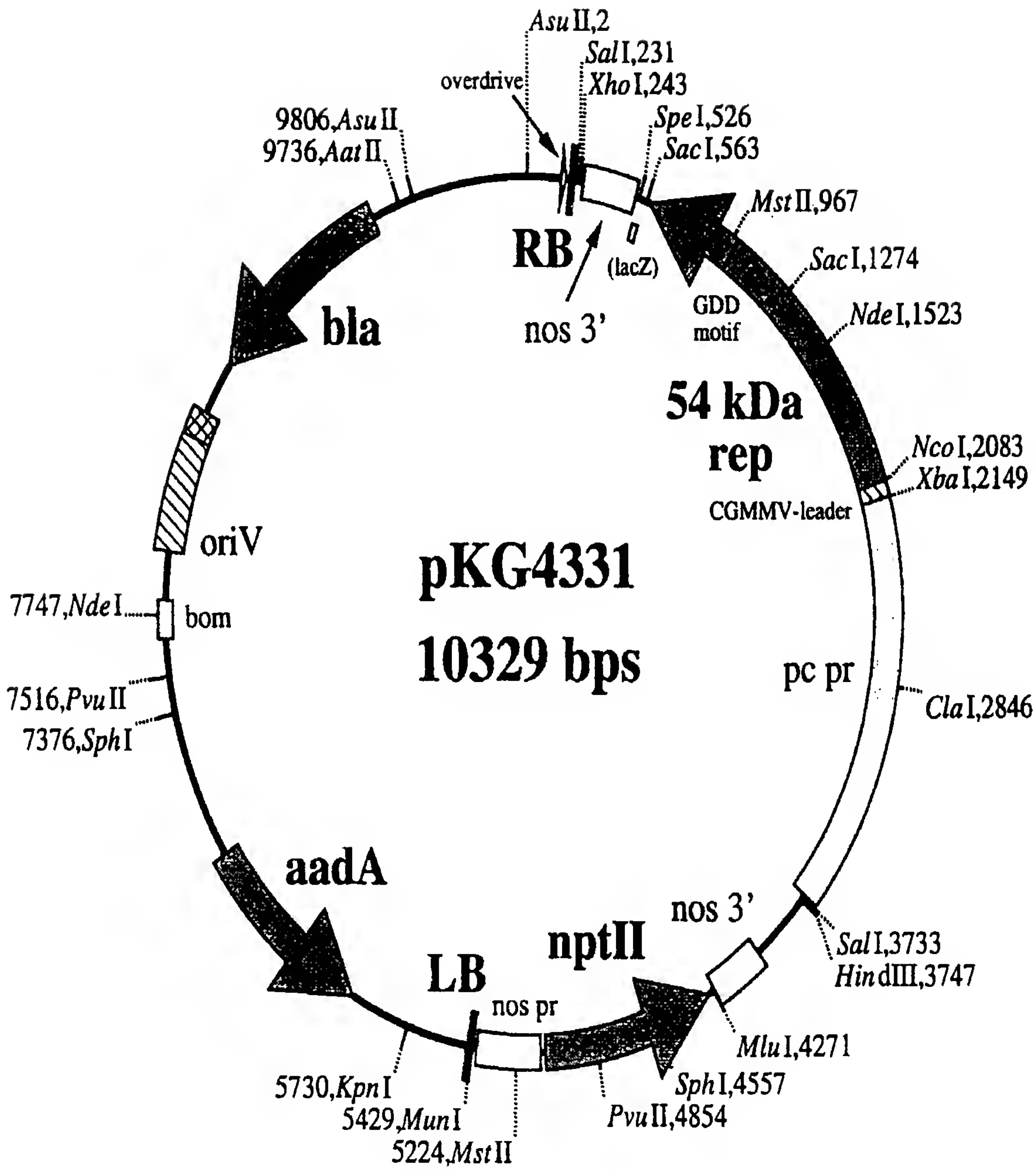

AsuII,2
SalI,231
XhoI,243
overdrive
SpeI,526
SacI,563
9806,AsuII
9736,AatII
MstII,967
RB
(lacZ)
SacI,1274
bla
nos 3'
GDD motif
NdeI,1523
54 kDa rep
NcoI,2083
XbaI,2149
CGMMV-leader
oriV
pKG4331
10329 bps
7747,NdeI
bom
pc pr
7516,PvuII
ClaI,2846
7376,SphI
aadA
nos 3'
nptII
SalI,3733
HindIII,3747
LB
nos pr
MluI,4271
SphI,4557
5730,KpnI
PvuII,4854
5429,MunI
5224,MstII

Fig 4

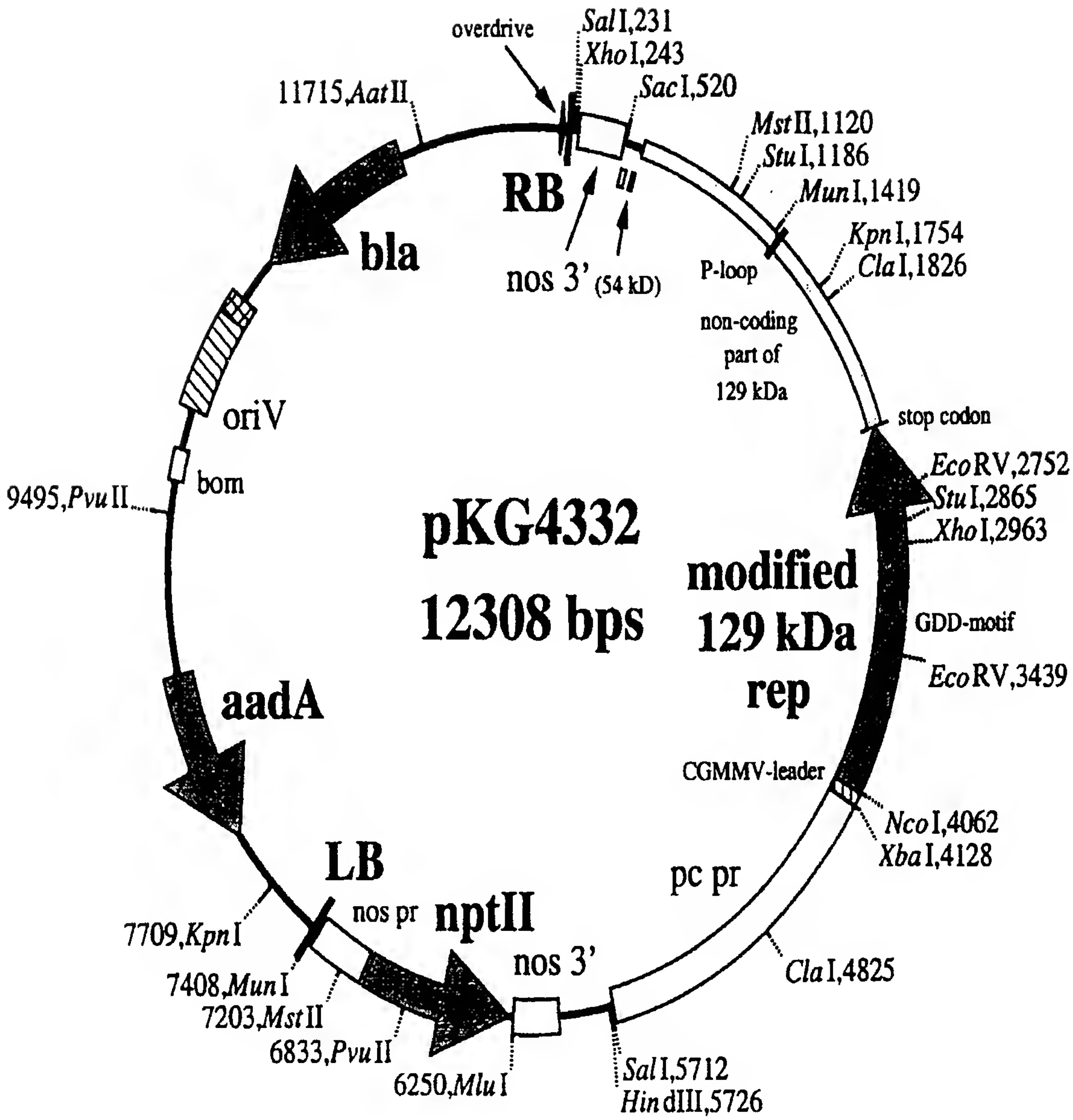

overdrive
SalI,231
XhoI,243
SacI,520
11715,AatII
MstII,1120
StuI,1186
MunI,1419
RB
KpnI,1754
ClaI,1826
bla
nos 3' (54 kD)
P-loop
non-coding
part of
129 kDa
oriV
stop codon
bom
EcoRV,2752
StuI,2865
9495,PvuII
pKG4332
12308 bps
XhoI,2963
modified
129 kDa
rep
GDD-motif
EcoRV,3439
aadA
CGMMV-leader
NcoI,4062
XbaI,4128
LB
pc pr
7709,KpnI
nos pr
nptII
ClaI,4825
nos 3'
7408,MunI
7203,MstII
6833,PvuII
6250,MluI
SalI,5712
HindIII,5726

Fig 5

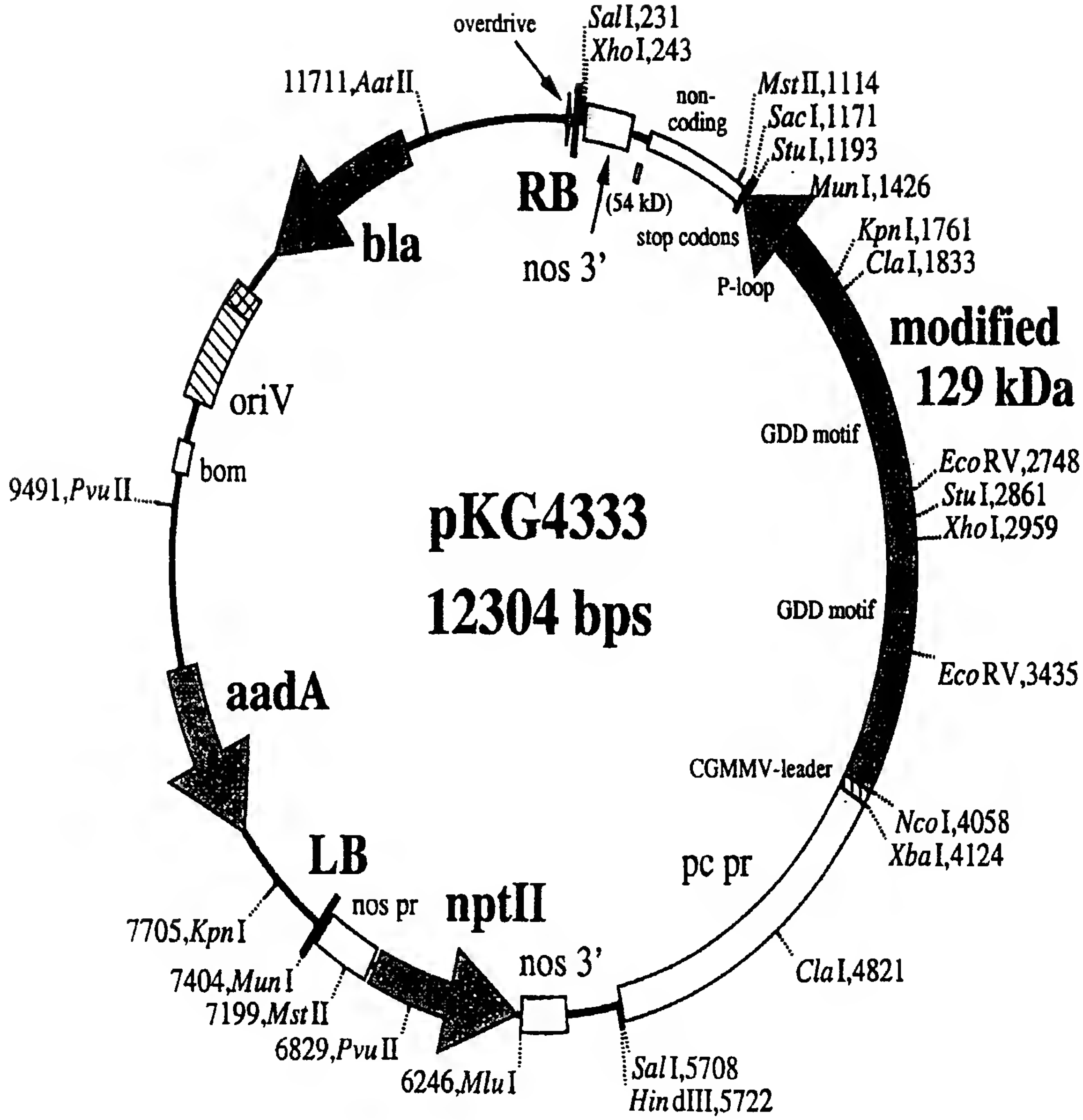

overdrive
SalI,231
XhoI,243
11711,AatII
non-coding
MstII,1114
SacI,1171
StuI,1193
MunI,1426
RB
(54 kD)
stop codons
KpnI,1761
ClaI,1833
bla
nos 3'
P-loop
modified
129 kDa
oriV
GDD motif
EcoRV,2748
StuI,2861
XhoI,2959
bom
9491,PvuII
pKG4333
12304 bps
GDD motif
EcoRV,3435
aadA
CGMMV-leader
NcoI,4058
XbaI,4124
LB
pc pr
nos pr
nptII
7705,KpnI
ClaI,4821
nos 3'
7404,MunI
7199,MstII
6829,PvuII
6246,MluI
SalI,5708
HindIII,5722

Fig 6

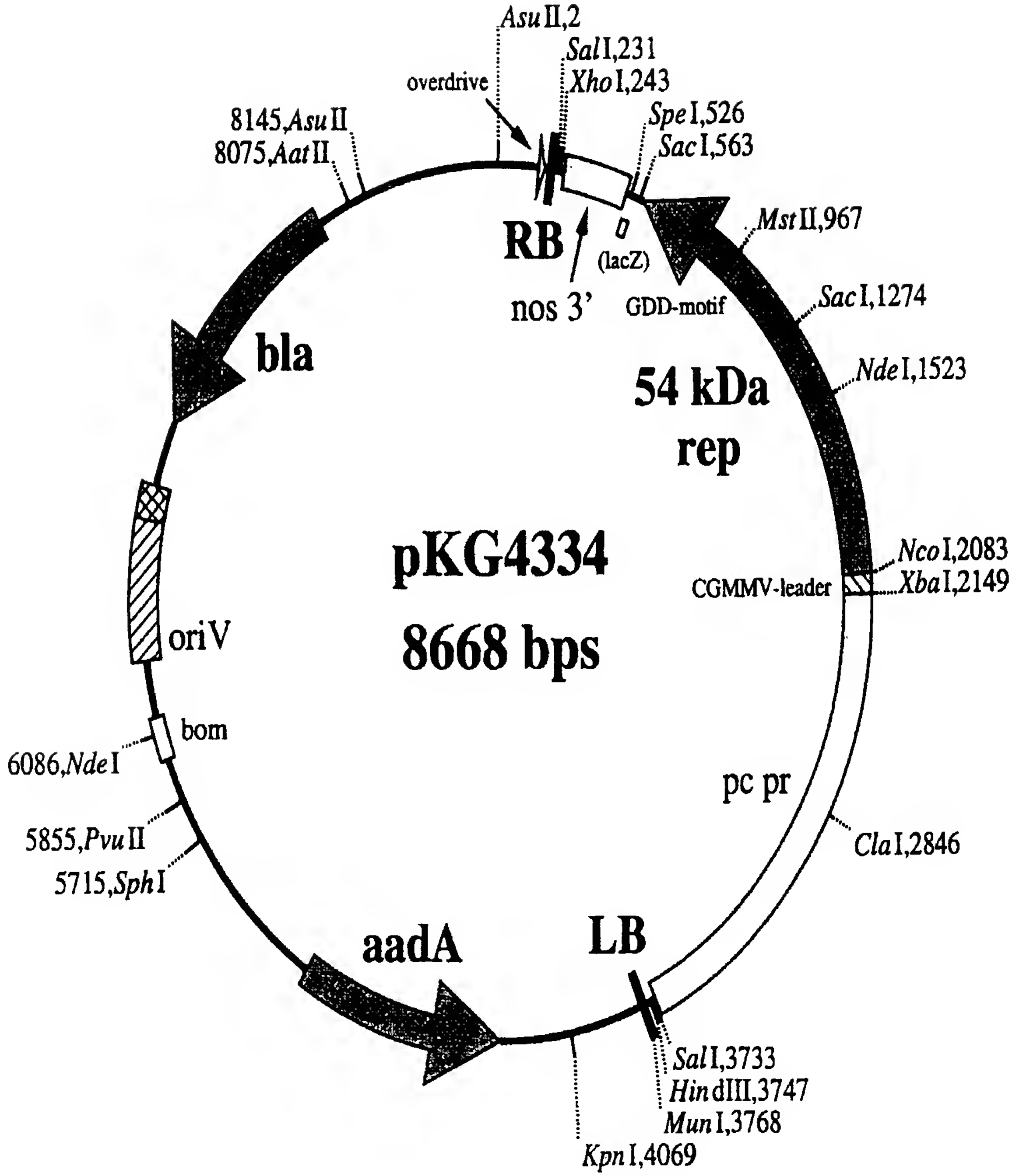

AsuII,2
SalI,231
XhoI,243
overdrive
SpeI,526
SacI,563
8145,AsuII
8075,AatII
MstII,967
RB
(lacZ)
nos 3'
GDD-motif
SacI,1274
bla
NdeI,1523
54 kDa
rep
pKG4334
8668 bps
NcoI,2083
CGMMV-leader
XbaI,2149
oriV
bom
6086,NdeI
pc pr
5855,PvuII
ClaI,2846
5715,SphI
aadA
LB
SalI,3733
HindIII,3747
MunI,3768
KpnI,4069

Fig 7

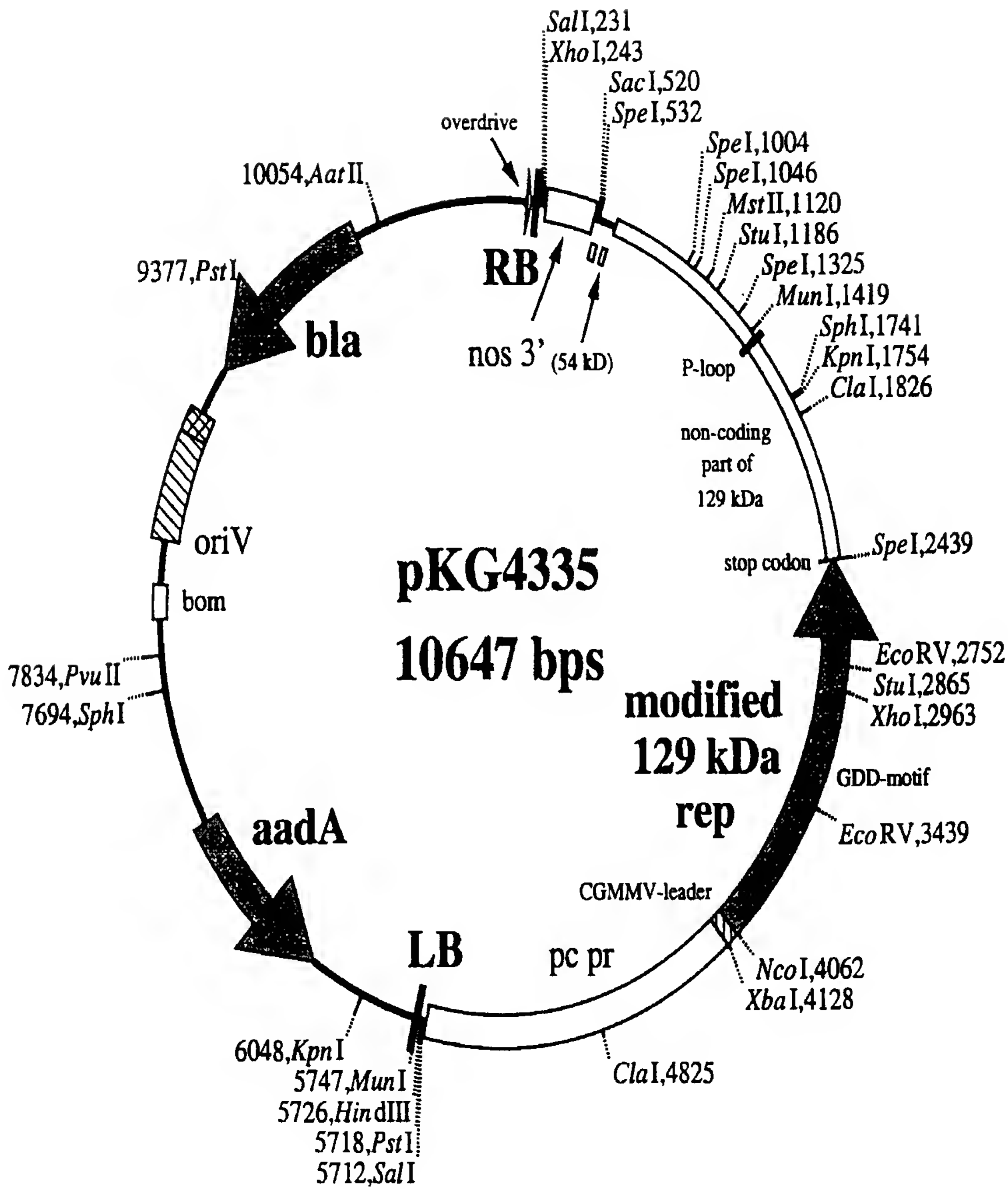
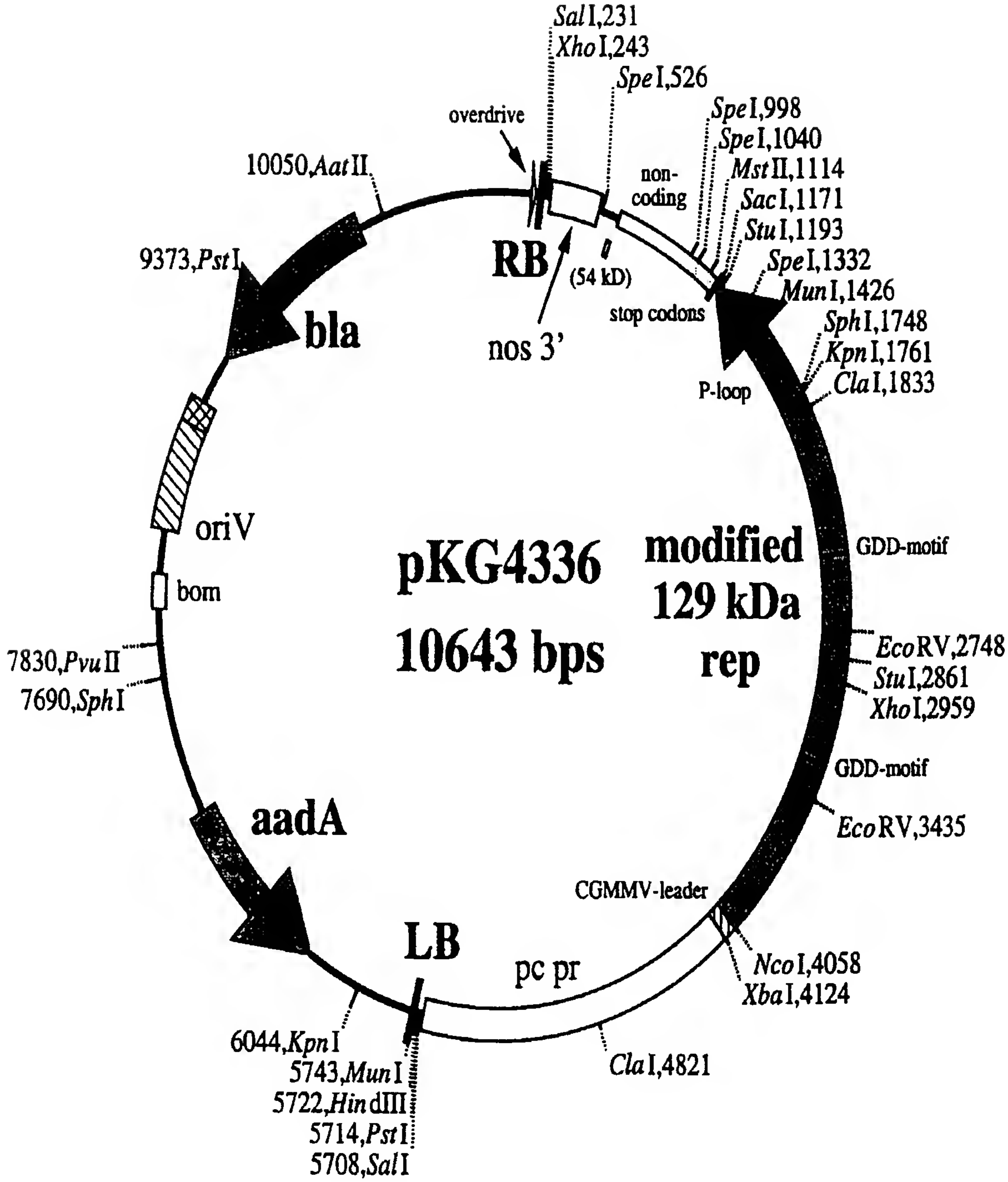

SalI,231
XhoI,243
SacI,520
SpeI,532
overdrive
SpeI,1004
SpeI,1046
MstII,1120
StuI,1186
SpeI,1325
MunI,1419
SphI,1741
KpnI,1754
ClaI,1826
10054,AatII
9377,PstI
RB
bla
nos 3'
(54 kD)
P-loop
non-coding
part of
129 kDa
oriV
stop codon
SpeI,2439
pKG4335
10647 bps
bom
EcoRV,2752
StuI,2865
XhoI,2963
7834,PvuII
7694,SphI
modified
129 kDa
rep
GDD-motif
EcoRV,3439
aadA
CGMMV-leader
LB
pc pr
NcoI,4062
XbaI,4128
6048,KpnI
5747,MunI
5726,HindIII
5718,PstI
5712,SalI
ClaI,4825

Fig 8


SalI,231
XhoI,243
SpeI,526
SpeI,998
SpeI,1040
MstII,1114
SacI,1171
StuI,1193
SpeI,1332
MunI,1426
SphI,1748
KpnI,1761
ClaI,1833
overdrive
non-coding
10050,AatII
9373,PstI
RB
(54 kD)
stop codons
nos 3'
bla
P-loop
oriV
bom
pKG4336
10643 bps
modified
129 kDa
rep
GDD-motif
EcoRV,2748
StuI,2861
XhoI,2959
7830,PvuII
7690,SphI
GDD-motif
EcoRV,3435
aadA
CGMMV-leader
LB
pc pr
NcoI,4058
XbaI,4124
6044,KpnI
5743,MunI
5722,HindIII
5714,PstI
5708,SalI
ClaI,4821

SEQUENCE LISTING

<110> Keygene N.V.

<120> Method for generating CGMMV resistance in plants, genetic constructs for use in said method, and CGMMV-resistant plants obtained via said method

<130> BO 42790

<140>
<141>

<160> 44

<170> PatentIn Ver. 2.1

<210> 1
<211> 3432
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to 129 kD replicase of CGMMV isolate 4.

<400> 1
atggcaaaca ttaatgaaca aatcaacaat caacgtgatg ctgctgctag cgggagaaat 60
aatctcgtta gtcagctagc atcaaagagg gtgtatgacg aggccgttcg ctcgttagat 120
catcaagata gacgcccaaa aatgaacttt tctcgtgtgg tcagtacaga gcacaccagg 180
cttgtcaccg atgcgtatcc ggagtttcg attagtttca ccgctaccaa gaattcagtt 240
cattcccttg cgggaggttt gaggcttctt gaattggaat acatgatgat gcaggtgcct 300
tatggttcac cttgctttga tattggcggt aattacacgc agcatttatt taaaggtaga 360
tcatatgtgc attgctgcaa tccgtgcctg gatcttaagg atgttgcgag gaatgtgatg 420
tacaacgaca tgatcacaca acatgtacag aggcacaaag gatctggtgg gtgtagacct 480
cttccgactt tccagataga tgctttcagg aggtatgaag attcgcccgt cgcagtcacc 540
tgtccagacg tttttcaaga atgctcctat gattttggga gtggtaggga taatcatgcg 600
gtttcattac attcgattta tgatatccct tattcttcga ttgggccagc tcttcatagg 660
aaaaacgtca gggtctgtta cgcagccttt catttctcgg aggcgttgct cctaggttcg 720
cccgtgggta atttaaatag tatagggct caatttaggg ttgatggtga cgatgtgcat 780
tttctttta gtgaggagtc aactttgcat tacactcata gtttggagaa tattaagttg 840
attgtaatgc gtacttattt ccctgctgat gataggttcg tgtatattaa ggagtttatg 900
gttaagcgtg tagacacttt tttttttagg ttagttaggg cagacacaca tatgctccat 960
aaatctgtag ggcactattc gaagtcgaaa tctgagtatt ttgcgttgaa cacccctccg 1020
attttccaag ataaggccac gtttctgtg tggtttcccg aagcgaagcg gaaggtgttg 1080
atacctaagt ttgaactctc gagatttctt tctggaaatg tgaaagtctc taggatgctt 1140
gtcgatgctg attttgtcca taccattatt aatcacatta gcacgtacga taacaaggcc 1200
ttagtgtgga agaatgtcca gtcttttgta gaatctatac gctctagggt aattgtaaac 1260
ggagtttccg taaaatctga atggaatgta ccggtcgatc agcttactga tatctcattc 1320
tcgatattcc ttctcgtgaa ggttagaaag gtgcagattg agttaatgtc tgataaggtt 1380
gtgatcgagg cgaggggttt gcttcggagg ttcgctgata gtctcaaatc cgccgtagaa 1440
ggactaggtg attgcgtcta tgatgctcta gttcaaaccg gttggtttga cacctctagc 1500
gacgaactga aagtattact acctgaaccg tttatgacct tttcagatta tctcgaaggg 1560
atgtacgagg cagatgcaaa aattgagaga gagagtgtct ctgagctgct tgcttccgga 1620
gatgatctgt tcaagaagat tgacgaaata aggaataatt acagcggagt tgaatttgat 1680
gtggagaaat ttcaagaatt ctgtaaagaa ctgaatgtta atcctatgct aatcggtcat 1740
gtgatcgaag ctattttttc acagaaggca ggggtaacag tcacgggcct aggcacgctc 1800
tctcctgaga tgggtgcttc cgttgcgtta tccaataatt ctgtagatac atgtgatgat 1860
atggacgtaa ctgaggatat ggaggaaata gtgttgatag cagacaagaa tcactcttat 1920
atttctccag aaatgtcgag atgggctagt atgaaatacg gcaataataa cggggcctta 1980

gttgagtaca aggtcggaac ctcgatgact ttacctgcca cctgggcaga aaagggtaag 2040
gctgttttac cgttgtcggg aatctgtgta agaaagcccc aattttcaaa gccactcgat 2100
gaggaggacg acttgaggtt atcaaacatg aatttcttta aggtgagtga tctgaagttg 2160
aagaagacta tcactccagt tgtttatact gggaccattc gagagaggca gatgaagaat 2220
tatatcgatt atctatcggc ttctctgggt tctacgcttg gtaatcttga gagaattgtt 2280
aggagtgact ggaatggtac cgaggagagc atgcaaactt ttggattgta cgattgcgag 2340
aagtgcaagt ggttactgtt gccatcggag aagaaacacg cctgggctgt agtcctggcg 2400
agtgatgata ccactcgtat aatctttctg tcgtatgacg aatccggttc tcctataatt 2460
gacaagaaaa attggaagcg gttcgctgtc tgttctgata ccaaagttta tagtgtaatt 2520
cgtagtttag aagtcttaaa taaggaggcc acagtcgatc ctggggtgta tataacttta 2580
gtcgatgggg ttccgggctg tggaaaaacc gctgaaatta tagcgagggt caattggaaa 2640
actgaccttg tgttgactcc cggaagggaa gcggctgcta tgatcaggcg aagagcctgt 2700
gccctacaca agtcacctgt agctactagt gataacgtta ggacttttga ttctttcgta 2760
atgaataaga aggtttttaa atttgacgcc gtctacgtag atgaaggtct tatggtccac 2820
acggggttgc tcaactttgc gttgaagatt tcgggttgta aaaaggcctt tgtcttcggt 2880
gatgctaagc aaattccgtt tattaataga gttatgaatt ttgattatcc taaggaatta 2940
agaactttga tagttgataa tgtagagcgt aggtatatta cccataggtg tcctagagat 3000
gtcactagtt ttcttaatac tatttataaa gctgcggttt ctaccactag tccggttgta 3060
cattccgtga aggcaataaa ggtttctggg gctggtattc tgaggcccga gttgacgaag 3120
atcaaaggga agatcataac gtttactcag tctgataaac aatccttgat caagagtggg 3180
tacaatgatg tgaatactgt gcatgagatt caggggggaga cctttgagga gacggcggtt 3240
gtgcgtgcaa caccgactcc aataggtctg attgcccgag attcaccaca cgtgttagtg 3300
gctttaacgc ggcacaccaa ggcaatggtg tattataccg ttgtgttcga tgccgtaaca 3360
agcataatag cggatgtgga aaaggtcgat cagtcgattt tgactatgtt tgctactact 3420
gtgcctacca aa 3432

<210> 2
<211> 1144
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: amino acid sequence corresponding to 129 kD replicase of CGMMV isolate 4

<400> 2
Met Ala Asn Ile Asn Glu Gln Ile Asn Asn Gln Arg Asp Ala Ala Ala
1 5 10 15

Ser Gly Arg Asn Asn Leu Val Ser Gln Leu Ala Ser Lys Arg Val Tyr
20 25 30

Asp Glu Ala Val Arg Ser Leu Asp His Gln Asp Arg Arg Pro Lys Met
35 40 45

Asn Phe Ser Arg Val Val Ser Thr Glu His Thr Arg Leu Val Thr Asp
50 55 60

Ala Tyr Pro Glu Phe Ser Ile Ser Phe Thr Ala Thr Lys Asn Ser Val
65 70 75 80

His Ser Leu Ala Gly Gly Leu Arg Leu Leu Glu Leu Glu Tyr Met Met
85 90 95

Met Gln Val Pro Tyr Gly Ser Pro Cys Phe Asp Ile Gly Gly Asn Tyr
100 105 110

Thr Gln His Leu Phe Lys Gly Arg Ser Tyr Val His Cys Cys Asn Pro
115 120 125

```
Cys Leu Asp Leu Lys Asp Val Ala Arg Asn Val Met Tyr Asn Asp Met
    130                 135                 140

Ile Thr Gln His Val Gln Arg His Lys Gly Ser Gly Gly Cys Arg Pro
145                 150                 155                 160

Leu Pro Thr Phe Gln Ile Asp Ala Phe Arg Arg Tyr Glu Asp Ser Pro
                165                 170                 175

Val Ala Val Thr Cys Pro Asp Val Phe Gln Glu Cys Ser Tyr Asp Phe
            180                 185                 190

Gly Ser Gly Arg Asp Asn His Ala Val Ser Leu His Ser Ile Tyr Asp
        195                 200                 205

Ile Pro Tyr Ser Ser Ile Gly Pro Ala Leu His Arg Lys Asn Val Arg
    210                 215                 220

Val Cys Tyr Ala Ala Phe His Phe Ser Glu Ala Leu Leu Leu Gly Ser
225                 230                 235                 240

Pro Val Gly Asn Leu Asn Ser Ile Gly Ala Gln Phe Arg Val Asp Gly
                245                 250                 255

Asp Asp Val His Phe Leu Phe Ser Glu Glu Ser Thr Leu His Tyr Thr
            260                 265                 270

His Ser Leu Glu Asn Ile Lys Leu Ile Val Met Arg Thr Tyr Phe Pro
        275                 280                 285

Ala Asp Asp Arg Phe Val Tyr Ile Lys Glu Phe Met Val Lys Arg Val
    290                 295                 300

Asp Thr Phe Phe Phe Arg Leu Val Arg Ala Asp Thr His Met Leu His
305                 310                 315                 320

Lys Ser Val Gly His Tyr Ser Lys Ser Lys Ser Glu Tyr Phe Ala Leu
                325                 330                 335

Asn Thr Pro Pro Ile Phe Gln Asp Lys Ala Thr Phe Ser Val Trp Phe
            340                 345                 350

Pro Glu Ala Lys Arg Lys Val Leu Ile Pro Lys Phe Glu Leu Ser Arg
        355                 360                 365

Phe Leu Ser Gly Asn Val Lys Val Ser Arg Met Leu Val Asp Ala Asp
    370                 375                 380

Phe Val His Thr Ile Ile Asn His Ile Ser Thr Tyr Asp Asn Lys Ala
385                 390                 395                 400

Leu Val Trp Lys Asn Val Gln Ser Phe Val Glu Ser Ile Arg Ser Arg
                405                 410                 415

Val Ile Val Asn Gly Val Ser Val Lys Ser Glu Trp Asn Val Pro Val
            420                 425                 430

Asp Gln Leu Thr Asp Ile Ser Phe Ser Ile Phe Leu Leu Val Lys Val
        435                 440                 445
```

```
Arg Lys Val Gln Ile Glu Leu Met Ser Asp Lys Val Val Ile Glu Ala
    450                 455                 460

Arg Gly Leu Leu Arg Arg Phe Ala Asp Ser Leu Lys Ser Ala Val Glu
465                 470                 475                 480

Gly Leu Gly Asp Cys Val Tyr Asp Ala Leu Val Gln Thr Gly Trp Phe
                485                 490                 495

Asp Thr Ser Ser Asp Glu Leu Lys Val Leu Leu Pro Glu Pro Phe Met
            500                 505                 510

Thr Phe Ser Asp Tyr Leu Glu Gly Met Tyr Glu Ala Asp Ala Lys Ile
        515                 520                 525

Glu Arg Glu Ser Val Ser Glu Leu Leu Ala Ser Gly Asp Asp Leu Phe
    530                 535                 540

Lys Lys Ile Asp Glu Ile Arg Asn Asn Tyr Ser Gly Val Glu Phe Asp
545                 550                 555                 560

Val Glu Lys Phe Gln Glu Phe Cys Lys Glu Leu Asn Val Asn Pro Met
                565                 570                 575

Leu Ile Gly His Val Ile Glu Ala Ile Phe Ser Gln Lys Ala Gly Val
            580                 585                 590

Thr Val Thr Gly Leu Gly Thr Leu Ser Pro Glu Met Gly Ala Ser Val
        595                 600                 605

Ala Leu Ser Asn Asn Ser Val Asp Thr Cys Asp Asp Met Asp Val Thr
    610                 615                 620

Glu Asp Met Glu Glu Ile Val Leu Ile Ala Asp Lys Asn His Ser Tyr
625                 630                 635                 640

Ile Ser Pro Glu Met Ser Arg Trp Ala Ser Met Lys Tyr Gly Asn Asn
                645                 650                 655

Asn Gly Ala Leu Val Glu Tyr Lys Val Gly Thr Ser Met Thr Leu Pro
            660                 665                 670

Ala Thr Trp Ala Glu Lys Gly Lys Ala Val Leu Pro Leu Ser Gly Ile
        675                 680                 685

Cys Val Arg Lys Pro Gln Phe Ser Lys Pro Leu Asp Glu Glu Asp Asp
    690                 695                 700

Leu Arg Leu Ser Asn Met Asn Phe Phe Lys Val Ser Asp Leu Lys Leu
705                 710                 715                 720

Lys Lys Thr Ile Thr Pro Val Val Tyr Thr Gly Thr Ile Arg Glu Arg
                725                 730                 735

Gln Met Lys Asn Tyr Ile Asp Tyr Leu Ser Ala Ser Leu Gly Ser Thr
            740                 745                 750

Leu Gly Asn Leu Glu Arg Ile Val Arg Ser Asp Trp Asn Gly Thr Glu
        755                 760                 765

Glu Ser Met Gln Thr Phe Gly Leu Tyr Asp Cys Glu Lys Cys Lys Trp
```

```
        770                 775                 780

Leu Leu Leu Pro Ser Glu Lys Lys His Ala Trp Ala Val Val Leu Ala
785                 790                 795                 800

Ser Asp Asp Thr Thr Arg Ile Ile Phe Leu Ser Tyr Asp Glu Ser Gly
                805                 810                 815

Ser Pro Ile Ile Asp Lys Lys Asn Trp Lys Arg Phe Ala Val Cys Ser
            820                 825                 830

Asp Thr Lys Val Tyr Ser Val Ile Arg Ser Leu Glu Val Leu Asn Lys
        835                 840                 845

Glu Ala Thr Val Asp Pro Gly Val Tyr Ile Thr Leu Val Asp Gly Val
    850                 855                 860

Pro Gly Cys Gly Lys Thr Ala Glu Ile Ile Ala Arg Val Asn Trp Lys
865                 870                 875                 . 880

Thr Asp Leu Val Leu Thr Pro Gly Arg Glu Ala Ala Ala Met Ile Arg
                885                 890                 895

Arg Arg Ala Cys Ala Leu His Lys Ser Pro Val Ala Thr Ser Asp Asn
            900                 905                 910

Val Arg Thr Phe Asp Ser Phe Val Met Asn Lys Lys Val Phe Lys Phe
        915                 920                 925

Asp Ala Val Tyr Val Asp Glu Gly Leu Met Val His Thr Gly Leu Leu
    930                 935                 940

Asn Phe Ala Leu Lys Ile Ser Gly Cys Lys Lys Ala Phe Val Phe Gly
945                 950                 955                 960

Asp Ala Lys Gln Ile Pro Phe Ile Asn Arg Val Met Asn Phe Asp Tyr
                965                 970                 975

Pro Lys Glu Leu Arg Thr Leu Ile Val Asp Asn Val Glu Arg Arg Tyr
            980                 985                 990

Ile Thr His Arg Cys Pro Arg Asp Val Thr Ser Phe Leu Asn Thr Ile
        995                 1000                1005

Tyr Lys Ala Ala Val Ser Thr Thr Ser Pro Val Val His Ser Val Lys
    1010                1015                1020

Ala Ile Lys Val Ser Gly Ala Gly Ile Leu Arg Pro Glu Leu Thr Lys
1025                1030                1035                1040

Ile Lys Gly Lys Ile Ile Thr Phe Thr Gln Ser Asp Lys Gln Ser Leu
                1045                1050                1055

Ile Lys Ser Gly Tyr Asn Asp Val Asn Thr Val His Glu Ile Gln Gly
            1060                1065                1070

Glu Thr Phe Glu Glu Thr Ala Val Val Arg Ala Thr Pro Thr Pro Ile
        1075                1080                1085

Gly Leu Ile Ala Arg Asp Ser Pro His Val Leu Val Ala Leu Thr Arg
    1090                1095                1100
```

```
His Thr Lys Ala Met Val Tyr Tyr Thr Val Val Phe Asp Ala Val Thr
1105                1110                1115                1120

Ser Ile Ile Ala Asp Val Glu Lys Val Asp Gln Ser Ile Leu Thr Met
            1125                1130                1135

Phe Ala Thr Thr Val Pro Thr Lys
        1140



<210> 3
<211> 1503
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide
      sequence corresponding to 57 kD protein of CGMMV
      isolate 4

<400> 3
atggagaatt cgctgtatgt ccaccgcaat atcttcctcc ctgttactaa gacagggttt 60
tatacggata tgcaggagtt ctatgacagg tgtcttccag ggaattcttt tgttctgaac 120
gatttcgatg ccgtcaccat gcggttgagg gataatgaat tcaatttgca accttgtaga 180
ttaactttaa gtaatttaga tccggtgccg gctttgatta agagtgaggc aaaagatttt 240
ctggttcccg tattgcgaac ggcttgcgaa aggccgcgta ttccgggtct tctcgaaaat 300
cttgttgcta tgataaagag gaatatgaat actcctgatt tggctgggac cgtggatata 360
actaatatgt ctatttctat agtagataat ttcttttctt cctttgtcag ggacgaggtt 420
ctacttgatc atttagattg cgttagagct agttctattc agagtttttc cgattggttt 480
tcttgtcagc caacctcggc ggttggccag ttagctaatt ttaacttcat agatttacct 540
gcctttgata cgtatatgca tatgattaaa aggcagccta agagtcggtt agatacttcg 600
attcagtccg aatatccggc cttacaaact attgtatatc atccgaaggt ggtaaacgca 660
gttttcgggc cggtttttaa gtatctgact actaagtttc ttagcatggt agataattct 720
aagtttttct tttatactag gaaaaagcca gaggatctgc aggaattttt ctcggatctt 780
tcttcccatt ctgattatga aattcttgag ctcgatgttt ctaaatatga taagtcgcag 840
tccgatttcc atttctctat cgagatggca atttgggaaa ggctgggact agatgatatt 900
ttagcttgga tgtggtctat gggtcataag agaactatac tgcaagattt ccaagctgga 960
ataaagacgc tcatttatta tcaaaggaag tctggcgacg taactacttt cataggtaat 1020
actttttatta ttgcagcgtg tgtagctagt atgttaccgt tagataagtg ttttaaggct 1080
agtttttgtg gtgatgattc gttaatctac cttcctaagg gtttggagta tcctgatatt 1140
caggctactg ccaatttggt ttggaatttt gaggcgaaac ttttccggaa gaagtatggt 1200
tactttctgcg ggaaatatat cattcatcac gccaacggtt gtattgttta ccctgacccct 1260
ttgaagttaa ttagtaaatt aggtagtaag agtcttgtag ggtacgagca tgtcgaggag 1320
tttcgtatat ctctcctcga tgtcgctcac agtttgttta atggtgctta tttccatttg 1380
ctcgacgatg caatccacga gttgtttcct aacgctgggg gttgcagttt tgtaataaat 1440
tgtttgtgta agtacttgag tgataagcgc cttttccgta gtctttatat agatgtctct 1500
aag                                                                1503


<210> 4
<211> 501
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: amino acid
      sequence corresponding to 57 kD protein of CGMMV
      isolate 4
```

```
<400> 4
Met Glu Asn Ser Leu Tyr Val His Arg Asn Ile Phe Leu Pro Val Thr
  1               5                  10                  15

Lys Thr Gly Phe Tyr Thr Asp Met Gln Glu Phe Tyr Asp Arg Cys Leu
            20                  25                  30

Pro Gly Asn Ser Phe Val Leu Asn Asp Phe Asp Ala Val Thr Met Arg
        35                  40                  45

Leu Arg Asp Asn Glu Phe Asn Leu Gln Pro Cys Arg Leu Thr Leu Ser
    50                  55                  60

Asn Leu Asp Pro Val Pro Ala Leu Ile Lys Ser Glu Ala Lys Asp Phe
65                  70                  75                  80

Leu Val Pro Val Leu Arg Thr Ala Cys Glu Arg Pro Arg Ile Pro Gly
                85                  90                  95

Leu Leu Glu Asn Leu Val Ala Met Ile Lys Arg Asn Met Asn Thr Pro
            100                 105                 110

Asp Leu Ala Gly Thr Val Asp Ile Thr Asn Met Ser Ile Ser Ile Val
        115                 120                 125

Asp Asn Phe Phe Ser Ser Phe Val Arg Asp Glu Val Leu Leu Asp His
    130                 135                 140

Leu Asp Cys Val Arg Ala Ser Ser Ile Gln Ser Phe Ser Asp Trp Phe
145                 150                 155                 160

Ser Cys Gln Pro Thr Ser Ala Val Gly Gln Leu Ala Asn Phe Asn Phe
                165                 170                 175

Ile Asp Leu Pro Ala Phe Asp Thr Tyr Met His Met Ile Lys Arg Gln
            180                 185                 190

Pro Lys Ser Arg Leu Asp Thr Ser Ile Gln Ser Glu Tyr Pro Ala Leu
        195                 200                 205

Gln Thr Ile Val Tyr His Pro Lys Val Val Asn Ala Val Phe Gly Pro
    210                 215                 220

Val Phe Lys Tyr Leu Thr Thr Lys Phe Leu Ser Met Val Asp Asn Ser
225                 230                 235                 240

Lys Phe Phe Phe Tyr Thr Arg Lys Lys Pro Glu Asp Leu Gln Glu Phe
                245                 250                 255

Phe Ser Asp Leu Ser Ser His Ser Asp Tyr Glu Ile Leu Glu Leu Asp
            260                 265                 270

Val Ser Lys Tyr Asp Lys Ser Gln Ser Asp Phe His Phe Ser Ile Glu
        275                 280                 285

Met Ala Ile Trp Glu Arg Leu Gly Leu Asp Asp Ile Leu Ala Trp Met
    290                 295                 300

Trp Ser Met Gly His Lys Arg Thr Ile Leu Gln Asp Phe Gln Ala Gly
305                 310                 315                 320
```

```
Ile Lys Thr Leu Ile Tyr Tyr Gln Arg Lys Ser Gly Asp Val Thr Thr
                325                 330                 335

Phe Ile Gly Asn Thr Phe Ile Ile Ala Ala Cys Val Ala Ser Met Leu
            340                 345                 350

Pro Leu Asp Lys Cys Phe Lys Ala Ser Phe Cys Gly Asp Asp Ser Leu
        355                 360                 365

Ile Tyr Leu Pro Lys Gly Leu Glu Tyr Pro Asp Ile Gln Ala Thr Ala
    370                 375                 380

Asn Leu Val Trp Asn Phe Glu Ala Lys Leu Phe Arg Lys Lys Tyr Gly
385                 390                 395                 400

Tyr Phe Cys Gly Lys Tyr Ile Ile His His Ala Asn Gly Cys Ile Val
                405                 410                 415

Tyr Pro Asp Pro Leu Lys Leu Ile Ser Lys Leu Gly Ser Lys Ser Leu
            420                 425                 430

Val Gly Tyr Glu His Val Glu Glu Phe Arg Ile Ser Leu Leu Asp Val
        435                 440                 445

Ala His Ser Leu Phe Asn Gly Ala Tyr Phe His Leu Leu Asp Asp Ala
    450                 455                 460

Ile His Glu Leu Phe Pro Asn Ala Gly Gly Cys Ser Phe Val Ile Asn
465                 470                 475                 480

Cys Leu Cys Lys Tyr Leu Ser Asp Lys Arg Leu Phe Arg Ser Leu Tyr
                485                 490                 495

Ile Asp Val Ser Lys
            500



<210> 5
<211> 4935
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide
      sequence corresponding to 186 kD protein of CGMMV
      isolate 4

<400> 5
atggcaaaca ttaatgaaca aatcaacaat caacgtgatg ctgctgctag cgggagaaat 60
aatctcgtta gtcagctagc atcaaagagg gtgtatgacg aggccgttcg ctcgttagat 120
catcaagata gacgcccaaa aatgaacttt tctcgtgtgg tcagtacaga gcacaccagg 180
cttgtcaccg atgcgtatcc ggagttttcg attagtttca ccgctaccaa gaattcagtt 240
cattcccttg cgggaggttt gaggcttctt gaattggaat acatgatgat gcaggtgcct 300
tatggttcac cttgctttga tattggcggt aattacacgc agcatttatt taaaggtaga 360
tcatatgtgc attgctgcaa tccgtgcctg gatcttaagg atgttgcgag gaatgtgatg 420
tacaacgaca tgatcacaca acatgtacag aggcacaaag gatctggtgg gtgtagacct 480
cttccgactt tccagataga tgctttcagg aggtatgaag attcgcccgt cgcagtcacc 540
tgtccagacg tttttcaaga atgctcctat gattttggga gtggtaggga taatcatgcg 600
gtttcattac attcgattta tgatatccct tattcttcga ttgggccagc tcttcatagg 660
aaaaacgtca gggtctgtta cgcagccttt catttctcgg aggcgttgct cctaggttcg 720
```

cccgtgggta atttaaatag tatagggct caatttaggg ttgatggtga cgatgtgcat 780
tttcttttta gtgaggagtc aactttgcat tacactcata gtttggagaa tattaagttg 840
attgtaatgc gtacttattt ccctgctgat gataggttcg tgtatattaa ggagtttatg 900
gttaagcgtg tagacacttt ttttttagg ttagttaggg cagacacaca tatgctccat 960
aaatctgtag ggcactattc gaagtcgaaa tctgagtatt ttgcgttgaa cacccctccg 1020
attttccaag ataaggccac gttttctgtg tggtttcccg aagcgaagcg gaaggtgttg 1080
atacctaagt ttgaactctc gagatttctt tctggaaatg tgaaagtctc taggatgctt 1140
gtcgatgctg attttgtcca taccattatt aatcacatta gcacgtacga taacaaggcc 1200
ttagtgtgga agaatgtcca gtcttttgta gaatctatac gctctagggt aattgtaaac 1260
ggagtttccg taaaatctga atggaatgta ccggtcgatc agcttactga tatctcattc 1320
tcgatattcc ttctcgtgaa ggttagaaag gtgcagattg agttaatgtc tgataaggtt 1380
gtgatcgagg cgaggggttt gcttcggagg ttcgctgata gtctcaaatc cgccgtagaa 1440
ggactaggtg attgcgtcta tgatgctcta gttcaaaccg gttggtttga cacctctagc 1500
gacgaactga aagtattact acctgaaccg tttatgacct tttcagatta tctcgaaggg 1560
atgtacgagg cagatgcaaa aattgagaga gagagtgtct ctgagctgct tgcttccgga 1620
gatgatctgt tcaagaagat tgacgaaata aggaataatt acagcggagt tgaatttgat 1680
gtggagaaat ttcaagaatt ctgtaaagaa ctgaatgtta atcctatgct aatcggtcat 1740
gtgatcgaag ctattttttc acagaaggca ggggtaacag tcacgggcct aggcacgctc 1800
tctcctgaga tgggtgcttc cgttgcgtta tccaataatt ctgtagatac atgtgatgat 1860
atggacgtaa ctgaggatat ggaggaaata gtgttgatag cagacaagaa tcactcttat 1920
atttctccag aaatgtcgag atgggctagt atgaaatacg gcaataataa cggggcctta 1980
gttgagtaca aggtcggaac ctcgatgact ttacctgcca cctgggcaga aaagggtaag 2040
gctgttttac cgttgtcggg aatctgtgta agaaagcccc aattttcaaa gccactcgat 2100
gaggaggacg acttgaggtt atcaaacatg aatttcttta aggtgagtga tctgaagttg 2160
aagaagacta tcactccagt tgtttatact gggaccattc gagagaggca gatgaagaat 2220
tatatcgatt atctatcggc ttctctgggt tctacgcttg gtaatcttga gagaattgtt 2280
aggagtgact ggaatggtac cgaggagagc atgcaaactt ttggattgta cgattgcgag 2340
aagtgcaagt ggttactgtt gccatcggag aagaaacacg cctgggctgt agtcctggcg 2400
agtgatgata ccactcgtat aatctttctg tcgtatgacg aatccggttc tcctataatt 2460
gacaagaaaa attggaagcg gttcgctgtc tgttctgata ccaaagttta tagtgtaatt 2520
cgtagtttag aagtcttaaa taaggaggcc acagtcgatc ctggggtgta tataacttta 2580
gtcgatgggg ttccgggctg tggaaaaacc gctgaaatta tagcgagggt caattggaaa 2640
actgaccttg tgttgactcc cggaagggaa gcggctgcta tgatcaggcg aagagcctgt 2700
gccctacaca agtcacctgt agctactagt gataacgtta ggacttttga ttctttcgta 2760
atgaataaga aggtttttaa atttgacgcc gtctacgtag atgaaggtct tatggtccac 2820
acggggttgc tcaactttgc gttgaagatt tcgggttgta aaaaggcctt tgtcttcggt 2880
gatgctaagc aaattccgtt tattaataga gttatgaatt ttgattatcc taaggaatta 2940
agaactttga tagttgataa tgtagagcgt aggtatatta cccataggtg tcctagagat 3000
gtcactagtt ttcttaatac tatttataaa gctgcggttt ctaccactag tccggttgta 3060
cattccgtga aggcaataaa ggtttctggg gctggtattc tgaggcccga gttgacgaag 3120
atcaaaggga agatcataac gtttactcag tctgataaac aatccttgat caagagtggg 3180
tacaatgatg tgaatactgt gcatgagatt caggggggaga cctttgagga gacggcggtt 3240
gtgcgtgcaa caccgactcc aataggtctg attgcccgag attcaccaca cgtgttagtg 3300
gctttaacgc ggcacaccaa ggcaatggtg tattataccg ttgtgttcga tgccgtaaca 3360
agcataatag cggatgtgga aaaggtcgat cagtcgattt tgactatgtt tgctactact 3420
gtgcctacca aaatggagaa ttcgctgtat gtccaccgca atatcttcct ccctgttact 3480
aagacagggt tttatacgga tatgcaggag ttctatgaca ggtgtcttcc agggaattct 3540
tttgttctga acgatttcga tgccgtcacc atgcggttga gggataatga attcaatttg 3600
caaccttgta gattaacttt aagtaattta gatccggtgc cggctttgat taagagtgag 3660
gcaaaagatt ttctggttcc cgtattgcga acggcttgcg aaaggccgcg tattccgggt 3720
cttctcgaaa atcttgttgc tatgataaag aggaatatga atactcctga tttggctggg 3780
accgtggata taactaatat gtctatttct atagtagata atttcttttc ttcctttgtc 3840
agggacgagg ttctacttga tcatttagat tgcgttagag ctagttctat tcagagtttt 3900
tccgattggt tttcttgtca gccaacctcg gcggttggcc agttagctaa ttttaacttc 3960
atagatttac ctgcctttga tacgtatatg catatgatta aaaggcagcc taagagtcgg 4020
ttagatactt cgattcagtc cgaatatccg gccttacaaa ctattgtata tcatccgaag 4080
gtggtaaacg cagttttcgg gccggttttt aagtatctga ctactaagtt tcttagcatg 4140
gtagataatt ctaagttttt cttttatact aggaaaaagc cagaggatct gcaggaattt 4200
ttctcggatc tttcttccca ttctgattat gaaattcttg agctcgatgt ttctaaatat 4260
gataagtcgc agtccgattt ccatttctct atcgagatgg caatttggga aaggctggga 4320
ctagatgata ttttagcttg gatgtggtct atgggtcata agagaactat actgcaagat 4380

```
ttccaagctg gaataaagac gctcatttat tatcaaagga agtctggcga cgtaactact 4440
ttcataggta atacttttat tattgcagcg tgtgtagcta gtatgttacc gttagataag 4500
tgttttaagg ctagtttttg tggtgatgat tcgttaatct accttcctaa gggtttggag 4560
tatcctgata ttcaggctac tgccaatttg gtttggaatt ttgaggcgaa acttttccgg 4620
aagaagtatg gttacttctg cgggaaatat atcattcatc acgccaacgg ttgtattgtt 4680
taccctgacc ctttgaagtt aattagtaaa ttaggtagta agagtcttgt agggtacgag 4740
catgtcgagg agtttcgtat atctctcctc gatgtcgctc acagtttgtt taatggtgct 4800
tatttccatt tgctcgacga tgcaatccac gagttgtttc ctaacgctgg gggttgcagt 4860
tttgtaataa attgtttgtg taagtacttg agtgataagc gccttttccg tagtctttat 4920
atagatgtct ctaag                                                  4935
```

<210> 6
<211> 1645
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: amino acid sequence corresponding to 186 kD protein of CGMMV isolate 4

<400> 6

```
Met Ala Asn Ile Asn Glu Gln Ile Asn Asn Gln Arg Asp Ala Ala Ala
 1               5                  10                  15

Ser Gly Arg Asn Asn Leu Val Ser Gln Leu Ala Ser Lys Arg Val Tyr
            20                  25                  30

Asp Glu Ala Val Arg Ser Leu Asp His Gln Asp Arg Arg Pro Lys Met
        35                  40                  45

Asn Phe Ser Arg Val Val Ser Thr Glu His Thr Arg Leu Val Thr Asp
    50                  55                  60

Ala Tyr Pro Glu Phe Ser Ile Ser Phe Thr Ala Thr Lys Asn Ser Val
65                  70                  75                  80

His Ser Leu Ala Gly Gly Leu Arg Leu Leu Glu Leu Glu Tyr Met Met
                85                  90                  95

Met Gln Val Pro Tyr Gly Ser Pro Cys Phe Asp Ile Gly Gly Asn Tyr
            100                 105                 110

Thr Gln His Leu Phe Lys Gly Arg Ser Tyr Val His Cys Cys Asn Pro
        115                 120                 125

Cys Leu Asp Leu Lys Asp Val Ala Arg Asn Val Met Tyr Asn Asp Met
    130                 135                 140

Ile Thr Gln His Val Gln Arg His Lys Gly Ser Gly Gly Cys Arg Pro
145                 150                 155                 160

Leu Pro Thr Phe Gln Ile Asp Ala Phe Arg Arg Tyr Glu Asp Ser Pro
                165                 170                 175

Val Ala Val Thr Cys Pro Asp Val Phe Gln Glu Cys Ser Tyr Asp Phe
            180                 185                 190

Gly Ser Gly Arg Asp Asn His Ala Val Ser Leu His Ser Ile Tyr Asp
        195                 200                 205
```

```
Ile Pro Tyr Ser Ser Ile Gly Pro Ala Leu His Arg Lys Asn Val Arg
    210                 215                 220

Val Cys Tyr Ala Ala Phe His Phe Ser Glu Ala Leu Leu Leu Gly Ser
225                 230                 235                     240

Pro Val Gly Asn Leu Asn Ser Ile Gly Ala Gln Phe Arg Val Asp Gly
                245                 250                     255

Asp Asp Val His Phe Leu Phe Ser Glu Glu Ser Thr Leu His Tyr Thr
            260                 265                 270

His Ser Leu Glu Asn Ile Lys Leu Ile Val Met Arg Thr Tyr Phe Pro
        275                 280                 285

Ala Asp Asp Arg Phe Val Tyr Ile Lys Glu Phe Met Val Lys Arg Val
    290                 295                 300

Asp Thr Phe Phe Phe Arg Leu Val Arg Ala Asp Thr His Met Leu His
305                 310                 315                     320

Lys Ser Val Gly His Tyr Ser Lys Ser Lys Ser Glu Tyr Phe Ala Leu
                325                 330                     335

Asn Thr Pro Pro Ile Phe Gln Asp Lys Ala Thr Phe Ser Val Trp Phe
            340                 345                 350

Pro Glu Ala Lys Arg Lys Val Leu Ile Pro Lys Phe Glu Leu Ser Arg
        355                 360                 365

Phe Leu Ser Gly Asn Val Lys Val Ser Arg Met Leu Val Asp Ala Asp
    370                 375                 380

Phe Val His Thr Ile Ile Asn His Ile Ser Thr Tyr Asp Asn Lys Ala
385                 390                 395                     400

Leu Val Trp Lys Asn Val Gln Ser Phe Val Glu Ser Ile Arg Ser Arg
                405                 410                     415

Val Ile Val Asn Gly Val Ser Val Lys Ser Glu Trp Asn Val Pro Val
            420                 425                 430

Asp Gln Leu Thr Asp Ile Ser Phe Ser Ile Phe Leu Leu Val Lys Val
        435                 440                 445

Arg Lys Val Gln Ile Glu Leu Met Ser Asp Lys Val Val Ile Glu Ala
    450                 455                 460

Arg Gly Leu Leu Arg Arg Phe Ala Asp Ser Leu Lys Ser Ala Val Glu
465                 470                 475                     480

Gly Leu Gly Asp Cys Val Tyr Asp Ala Leu Val Gln Thr Gly Trp Phe
                485                 490                     495

Asp Thr Ser Ser Asp Glu Leu Lys Val Leu Leu Pro Glu Pro Phe Met
            500                 505                 510

Thr Phe Ser Asp Tyr Leu Glu Gly Met Tyr Glu Ala Asp Ala Lys Ile
        515                 520                 525
```

```
Glu Arg Glu Ser Val Ser Glu Leu Leu Ala Ser Gly Asp Asp Leu Phe
    530                 535                 540

Lys Lys Ile Asp Glu Ile Arg Asn Asn Tyr Ser Gly Val Glu Phe Asp
545                 550                 555                 560

Val Glu Lys Phe Gln Glu Phe Cys Lys Glu Leu Asn Val Asn Pro Met
                565                 570                 575

Leu Ile Gly His Val Ile Glu Ala Ile Phe Ser Gln Lys Ala Gly Val
            580                 585                 590

Thr Val Thr Gly Leu Gly Thr Leu Ser Pro Glu Met Gly Ala Ser Val
        595                 600                 605

Ala Leu Ser Asn Asn Ser Val Asp Thr Cys Asp Asp Met Asp Val Thr
    610                 615                 620

Glu Asp Met Glu Glu Ile Val Leu Ile Ala Asp Lys Asn His Ser Tyr
625                 630                 635                 640

Ile Ser Pro Glu Met Ser Arg Trp Ala Ser Met Lys Tyr Gly Asn Asn
                645                 650                 655

Asn Gly Ala Leu Val Glu Tyr Lys Val Gly Thr Ser Met Thr Leu Pro
            660                 665                 670

Ala Thr Trp Ala Glu Lys Gly Lys Ala Val Leu Pro Leu Ser Gly Ile
        675                 680                 685

Cys Val Arg Lys Pro Gln Phe Ser Lys Pro Leu Asp Glu Glu Asp Asp
    690                 695                 700

Leu Arg Leu Ser Asn Met Asn Phe Phe Lys Val Ser Asp Leu Lys Leu
705                 710                 715                 720

Lys Lys Thr Ile Thr Pro Val Val Tyr Thr Gly Thr Ile Arg Glu Arg
                725                 730                 735

Gln Met Lys Asn Tyr Ile Asp Tyr Leu Ser Ala Ser Leu Gly Ser Thr
            740                 745                 750

Leu Gly Asn Leu Glu Arg Ile Val Arg Ser Asp Trp Asn Gly Thr Glu
        755                 760                 765

Glu Ser Met Gln Thr Phe Gly Leu Tyr Asp Cys Glu Lys Cys Lys Trp
    770                 775                 780

Leu Leu Leu Pro Ser Glu Lys Lys His Ala Trp Ala Val Val Leu Ala
785                 790                 795                 800

Ser Asp Asp Thr Thr Arg Ile Ile Phe Leu Ser Tyr Asp Glu Ser Gly
                805                 810                 815

Ser Pro Ile Ile Asp Lys Lys Asn Trp Lys Arg Phe Ala Val Cys Ser
            820                 825                 830

Asp Thr Lys Val Tyr Ser Val Ile Arg Ser Leu Glu Val Leu Asn Lys
        835                 840                 845

Glu Ala Thr Val Asp Pro Gly Val Tyr Ile Thr Leu Val Asp Gly Val
```

```
    850                 855                 860

Pro Gly Cys Gly Lys Thr Ala Glu Ile Ile Ala Arg Val Asn Trp Lys
865                 870                 875                 880

Thr Asp Leu Val Leu Thr Pro Gly Arg Glu Ala Ala Ala Met Ile Arg
                885                 890                 895

Arg Arg Ala Cys Ala Leu His Lys Ser Pro Val Ala Thr Ser Asp Asn
            900                 905                 910

Val Arg Thr Phe Asp Ser Phe Val Met Asn Lys Lys Val Phe Lys Phe
        915                 920                 925

Asp Ala Val Tyr Val Asp Glu Gly Leu Met Val His Thr Gly Leu Leu
    930                 935                 940

Asn Phe Ala Leu Lys Ile Ser Gly Cys Lys Lys Ala Phe Val Phe Gly
945                 950                 955               . 960

Asp Ala Lys Gln Ile Pro Phe Ile Asn Arg Val Met Asn Phe Asp Tyr
                965                 970                 975

Pro Lys Glu Leu Arg Thr Leu Ile Val Asp Asn Val Glu Arg Arg Tyr
            980                 985                 990

Ile Thr His Arg Cys Pro Arg Asp Val Thr Ser Phe Leu Asn Thr Ile
        995                 1000                1005

Tyr Lys Ala Ala Val Ser Thr Thr Ser Pro Val Val His Ser Val Lys
   1010                 1015                1020

Ala Ile Lys Val Ser Gly Ala Gly Ile Leu Arg Pro Glu Leu Thr Lys
1025                1030                1035                1040

Ile Lys Gly Lys Ile Ile Thr Phe Thr Gln Ser Asp Lys Gln Ser Leu
               1045                1050                1055

Ile Lys Ser Gly Tyr Asn Asp Val Asn Thr Val His Glu Ile Gln Gly
           1060                1065                1070

Glu Thr Phe Glu Glu Thr Ala Val Val Arg Ala Thr Pro Thr Pro Ile
       1075                1080                1085

Gly Leu Ile Ala Arg Asp Ser Pro His Val Leu Val Ala Leu Thr Arg
   1090                1095                1100

His Thr Lys Ala Met Val Tyr Tyr Thr Val Val Phe Asp Ala Val Thr
1105                1110                1115                1120

Ser Ile Ile Ala Asp Val Glu Lys Val Asp Gln Ser Ile Leu Thr Met
               1125                1130                1135

Phe Ala Thr Thr Val Pro Thr Lys Met Glu Asn Ser Leu Tyr Val His
           1140                1145                1150

Arg Asn Ile Phe Leu Pro Val Thr Lys Thr Gly Phe Tyr Thr Asp Met
       1155                1160                1165

Gln Glu Phe Tyr Asp Arg Cys Leu Pro Gly Asn Ser Phe Val Leu Asn
   1170                1175                1180
```

Asp Phe Asp Ala Val Thr Met Arg Leu Arg Asp Asn Glu Phe Asn Leu
1185 1190 1195 1200

Gln Pro Cys Arg Leu Thr Leu Ser Asn Leu Asp Pro Val Pro Ala Leu
1205 1210 1215

Ile Lys Ser Glu Ala Lys Asp Phe Leu Val Pro Val Leu Arg Thr Ala
1220 1225 1230

Cys Glu Arg Pro Arg Ile Pro Gly Leu Leu Glu Asn Leu Val Ala Met
1235 1240 1245

Ile Lys Arg Asn Met Asn Thr Pro Asp Leu Ala Gly Thr Val Asp Ile
1250 1255 1260

Thr Asn Met Ser Ile Ser Ile Val Asp Asn Phe Phe Ser Ser Phe Val
1265 1270 1275 1280

Arg Asp Glu Val Leu Leu Asp His Leu Asp Cys Val Arg Ala Ser Ser
1285 1290 1295

Ile Gln Ser Phe Ser Asp Trp Phe Ser Cys Gln Pro Thr Ser Ala Val
1300 1305 1310

Gly Gln Leu Ala Asn Phe Asn Phe Ile Asp Leu Pro Ala Phe Asp Thr
1315 1320 1325

Tyr Met His Met Ile Lys Arg Gln Pro Lys Ser Arg Leu Asp Thr Ser
1330 1335 1340

Ile Gln Ser Glu Tyr Pro Ala Leu Gln Thr Ile Val Tyr His Pro Lys
1345 1350 1355 1360

Val Val Asn Ala Val Phe Gly Pro Val Phe Lys Tyr Leu Thr Thr Lys
1365 1370 1375

Phe Leu Ser Met Val Asp Asn Ser Lys Phe Phe Phe Tyr Thr Arg Lys
1380 1385 1390

Lys Pro Glu Asp Leu Gln Glu Phe Phe Ser Asp Leu Ser Ser His Ser
1395 1400 1405

Asp Tyr Glu Ile Leu Glu Leu Asp Val Ser Lys Tyr Asp Lys Ser Gln
1410 1415 1420

Ser Asp Phe His Phe Ser Ile Glu Met Ala Ile Trp Glu Arg Leu Gly
1425 1430 1435 1440

Leu Asp Asp Ile Leu Ala Trp Met Trp Ser Met Gly His Lys Arg Thr
1445 1450 1455

Ile Leu Gln Asp Phe Gln Ala Gly Ile Lys Thr Leu Ile Tyr Tyr Gln
1460 1465 1470

Arg Lys Ser Gly Asp Val Thr Thr Phe Ile Gly Asn Thr Phe Ile Ile
1475 1480 1485

Ala Ala Cys Val Ala Ser Met Leu Pro Leu Asp Lys Cys Phe Lys Ala
1490 1495 1500

```
Ser Phe Cys Gly Asp Asp Ser Leu Ile Tyr Leu Pro Lys Gly Leu Glu
1505                1510                1515                1520

Tyr Pro Asp Ile Gln Ala Thr Ala Asn Leu Val Trp Asn Phe Glu Ala
                1525                1530                1535

Lys Leu Phe Arg Lys Lys Tyr Gly Tyr Phe Cys Gly Lys Tyr Ile Ile
            1540                1545                1550

His His Ala Asn Gly Cys Ile Val Tyr Pro Asp Pro Leu Lys Leu Ile
        1555                1560                1565

Ser Lys Leu Gly Ser Lys Ser Leu Val Gly Tyr Glu His Val Glu Glu
    1570                1575                1580

Phe Arg Ile Ser Leu Leu Asp Val Ala His Ser Leu Phe Asn Gly Ala
1585                1590                1595                1600

Tyr Phe His Leu Leu Asp Asp Ala Ile His Glu Leu Phe Pro Asn Ala
                1605                1610                1615

Gly Gly Cys Ser Phe Val Ile Asn Cys Leu Cys Lys Tyr Leu Ser Asp
            1620                1625                1630

Lys Arg Leu Phe Arg Ser Leu Tyr Ile Asp Val Ser Lys
        1635                1640                1645


<210> 7
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide
      sequence corresponding to sequence encoding coat
      protein of CGMMV isolate 1

<400> 7
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtaatggc gagagacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccagttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgctc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttc caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtt gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ccttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gaatcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg  1139
```

<210> 8
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 2

<400> 8
```
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtaatggc gagagacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccagttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgctc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttc caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtt gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gaatcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg  1139
```

<210> 9
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 3

<400> 9
```
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtaatggc gagagacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggtattc taaacctcca aatcggaggt 480
cgggctctgc ttctgaagag tccagttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgttc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttc caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtc gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gtttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gagtcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaaaggg tttgatgttt atgatagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
```

tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg 1139

<210> 10
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 4

<400> 10
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctccgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtttagt tgccactact aactctatta ttaaaagggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcag ctgtaatggc gagagacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccggttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgttc ccgttagaac tctacttaat 600
tttctggtgg cgtcgcaagg tactgctttc caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtc gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gagtcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg 1139

<210> 11
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 5

<400> 11
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctccgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtttagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcag ctgtaatggc gagagacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccggttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgttc ccgttagaac tctacttaat 600
tttctggtgg cgtcgcaagg tactgctttc caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtc gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gagtcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900

gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg 1139

<210> 12
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 6

<400> 12
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaagga gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtaatggc gagagacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccagttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgttc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttc caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtt gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gaatcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgacagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg 1139

<210> 13
<211> 1138
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 7

<400> 13
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtgatggc gagggacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccagttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgttc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttc caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtt gtagatatta attctaggtt cccgaatgcg 720

```
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgttgatc ctagcaatcc gacgactgct 840
gagtcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa ctctcgcgtg tgacgttgaa gtttctatgg gcaagccg   1138
```

<210> 14
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 8

<400> 14

```
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggttg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggtattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttatctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtaatggc gagagacccc agttataggt 420
ctaggtcgca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccagttctg tttctttcga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcttatgctc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttc caaatccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtt gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gaatcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcttttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg  1139
```

<210> 15
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to sequence encoding coat protein of CGMMV isolate 9

<400> 15

```
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggctg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggcattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttgtctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtaatggc tagagacccc agttataggt 420
ctaggtcaca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccagttctg tttcttttga agatggctta caatccgatc 540
```

```
acgcctagca aacttattgc gtttagtgct tcatatgttc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttt caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtt gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gagtcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcctttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgttttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg  1139
```

```
<210> 16
<211> 1139
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide
      sequence corresponding to sequence encoding coat
      protein of CGMMV isolate 10.

<400> 16
aattcggctt ctgtaggggt ggtgctactg ttgctttggt tgacacaagg atgcattctg 60
ttgcagaagg aactatatgc aaattttcag ctcccgccac cgtccgcgag ttctctgtta 120
ggttcatccc taactattct gtcgtggctg cggatgccct tcgcgatcct tggtctttat 180
ttgtgaggct ctctaacgta ggcattaagg atggttttca tccattaact ttagaggtcg 240
cctgtctagt tgccactact aactctatta ttaaaaaggg gcttagagct tctgtagttg 300
agtccgttgt ctcttccgat cagtcgattg ttctagattc tttgtctgag aaagttgagc 360
ctttcttcga taaagtccct atttcagcgg ctgtaatggc tagagacccc agttataggt 420
ctaggtcaca gtctgtcgtt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
cggactctgc ttctgaagag tccagttctg tttcttttga agatggctta caatccgatc 540
acgcctagca aacttattgc gtttagtgct tcatatgttc ccgttagaac tttacttaat 600
tttctagtgg cgtcgcaagg tactgctttt caaacccagg caggaagaga ttccttccgt 660
gagtctttgt ctgcgttacc ttcatccgtt gtagatatta attctaggtt cccgagtgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttcttagc 780
tctacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc gacgactgct 840
gagtcgctta acgcagttaa gcgtactgac gatgcgtcta cagccgctag ggctgagata 900
gataatttaa tagaatcaat ctctaagggg tttgatgttt atgatagggc ttcctttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcca aggcttagcc ttgagggtct 1020
tctgacggtg gtgcacacca tagtgcatag tgctttcccg ttcactttaa tcgaacggtt 1080
tgctcattgg tttgcgaaaa cctctcgcgt gtgacgttga agtttctatg ggcaagccg  1139
```

```
<210> 17
<211> 3429
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide
      sequence corresponding to 129 kD replicase of
      CGMMV strain SH

<400> 17
atggcaaaca ttaatgaaca aatcaacaac caacgtgacg ccgcggctag cgggagaaac 60
aatctcgtta gccaattggc gtcaaaaagg gtgtatgacg aggctgttcg ctcgttggat 120
catcaagaca gacgcccgaa aatgaatttt tctcgtgtgg tcagcacaga gcacaccagg 180
cttgtaactg acgcgtatcc ggagttttcg attagcttta ccgccaccaa gaactctgta 240
cactcccttg cgggtggtct gaggcttctt gaattggaat atatgatgat gcaggtgccc 300
tacggctcac cttgttatga catcggcggt aactatacgc agcacttgtt caaaggtaga 360
```

```
tcatatgtgc attgctgcaa tccgtgccta gatcttaagg atgttgcgag gaatgtgatg 420
tacaacgata tgattacgca acatgtacag aggcacaagg gatcttgcgg gtgcagacct 480
cttccaactt tccagataga tgcattcagg aggtacgata gttctccctg tgcggtcacc 540
tgttcagacg ttttccaaga gtgttcctat gattttggga gtggtaggga taatcatgca 600
gtctcgttgc attcaatcta cgatatccct tattcttcga tcggacctgc tcttcatagg 660
aagaatgtgc gagtttgtta tgcagccttt catttctcgg aggcattgct tttaggttcg 720
cctgtaggta atttaaatag tattggggct cagtttaggg tcgatggtga tgatgtgcat 780
tttcttttta gtgaagagtc tactttgcat tatactcata gtttagaaaa tatcaagtta 840
atcgtgatgc gtacttactt tcctgctgat gataggtttg tatatattaa ggagttcatg 900
gttaagcgtg tggatacttt tttctttagg ttggtcagag cagatacaca catgcttcat 960
aaatctgtgg ggcactattc gaaatcgaag tctgagtact tcgcgctgaa tacccctccg 1020
atcttccaag ataaagccac gtttctgtg tggtttcctg aagcgaagaa ggtgttgata 1080
cccaagtttg aactttcgag attcctttct gggaatgtga aaatctctag gatgcttgtc 1140
gatgctgatt tcgtccatac cattattaat cacattagca cgtatgataa caaggcctta 1200
gtgtggaaga atgttcagtc ctttgtggaa tccatacgtt caagagtaat tgtaaacgga 1260
gtttccgtga aatctgagtg gaacgtaccg gttgatcagc tcactgatat ctcgttctcg 1320
atattccctc tcgtgaaggt taggaaggta cagatcgagt taatgtctga taaagttgta 1380
atcgaggcga ggggtttgct tcggaggttc gcagacagtc ttaaatctgc cgtagaagga 1440
ctaggtgatt gcgtctatga tgctctagtt caaaccggct ggtttgacac ctctagcgac 1500
gaactgaaag tattgctacc tgaaccgttt atgacctttt cggattatct tgaagggatg 1560
tacgaggcag atgcaaagat cgagagagag agtgtctctg agttgctcgc ttccggtgat 1620
gatttgttca agaaaatcga tgagataaga aacaattaca gtggagtcga atttgatgta 1680
gagaaattcc aagaattttg caaggaactg aatgttaatc ctatgctaat tggccatgtc 1740
atcgaagcta ttttttcgca gaaggctggg gtaacagtaa cgggtctggg cacgctctct 1800
cctgagatgg gcgcttctgt tgcgttatcc agtacctctg tagatacatg tgaagatatg 1860
gatgtaactg aagatatgga ggatatagtg ttgatggcgg acaagagtca ttcttacatg 1920
tcccctgaaa tggcgagatg ggctgatgtt aaatatggca acaataaagg ggctctagtc 1980
gagtacaaag tcggaacctc gatgacttta cctgccacct gggcagagaa agttaaggct 2040
gtcttaccgt tgtcggggat ctgtgtgagg aaaccccaat tttcgaagcc gcttgatgag 2100
gaagatgact tgaggttatc aaacatgaat ttctttaagg tgagcgatct aaagttgaag 2160
aagactatca ctccagtcgt ttacactggg accattcgag agaggcaaat gaagaattat 2220
attgattact tatcggcctc tcttggttcc acgctgggta atctggagag aatcgtgcgg 2280
agtgattgga atggtactga ggagagtatg caaacgttcg ggttgtatga ctgcgaaaag 2340
tgcaagtggt tattgttgcc agccgagaag aagcacgcat gggccgtggt tctggcaagt 2400
gacgatacca ctcgcataat cttcctttca tatgacgaat ctggttctcc tataattgat 2460
aagaaaaact ggaagcgatt tgctgtctgt tccgagacca aagtctatag tgtaattcgt 2520
agcttagagg ttctaaataa ggaagcaata gtcgaccccg ggttcacat aacattagtt 2580
gacggagtgc cgggttgtgg aaagaccgcc gagattatag cgagggtcaa ttggaaaact 2640
gatctagtat tgactcccgg aagggaggca gctgctatga ttaggcggag agcctgcgcc 2700
ctgcacaagt cacctgtggc aaccaatgac aacgtcagaa ctttcgattc ttttgtgatg 2760
aataggaaaa tcttcaagtt tgacgctgtc tatgttgacg agggtctgat ggtccatacg 2820
ggattactta attttgcgtt aaagatctca ggttgtaaaa aagccttcgt ctttggtgat 2880
gctaagcaaa tcccgtttat aaacagagtc atgaatttcg attatcctaa ggagttaaga 2940
actttaatag tcgataatgt agagcgtagg tatgtcaccc ataggtgtcc tagagatgtc 3000
actagttttc ttaatactat ctataaagcc gctgtcgcta ctactagtcc ggttgtacat 3060
tctgtgaagg caattaaagt gtcaggggcc ggtattctga ggcctgagtt gacaaagatc 3120
aaaggaaaga taataacgtt tactcaatct gataagcagt ccttgatcaa gagtgggtac 3180
aatgatgtga atactgtgca tgaaattcag ggagaaacct ttgaggagac ggcagttgtg 3240
cgtgccaccc cgactccaat aggtttgatt gcccgtgatt caccacatgt actagtggcc 3300
ttaactaggc acactaaggc aatggtgtat tatactgttg tattcgatgc agttacaagt 3360
ataatagcgg atgtggaaaa ggtcgatcag tcgatcttga ccatgtttgc taccactgtg 3420
cctaccaaa                                                         3429
```

<210> 18
<211> 1143
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: amino acid

sequence corresponding to 129 kD replicase of CGMMV strain SH

<400> 18

```
Met Ala Asn Ile Asn Glu Gln Ile Asn Asn Gln Arg Asp Ala Ala Ala
1               5                   10                  15

Ser Gly Arg Asn Asn Leu Val Ser Gln Leu Ala Ser Lys Arg Val Tyr
            20                  25                  30

Asp Glu Ala Val Arg Ser Leu Asp His Gln Asp Arg Arg Pro Lys Met
        35                  40                  45

Asn Phe Ser Arg Val Val Ser Thr Glu His Thr Arg Leu Val Thr Asp
    50                  55                  60

Ala Tyr Pro Glu Phe Ser Ile Ser Phe Thr Ala Thr Lys Asn Ser Val
65                  70                  75                  80

His Ser Leu Ala Gly Gly Leu Arg Leu Leu Glu Leu Glu Tyr Met Met
                85                  90                  95

Met Gln Val Pro Tyr Gly Ser Pro Cys Tyr Asp Ile Gly Gly Asn Tyr
            100                 105                 110

Thr Gln His Leu Phe Lys Gly Arg Ser Tyr Val His Cys Cys Asn Pro
        115                 120                 125

Cys Leu Asp Leu Lys Asp Val Ala Arg Asn Val Met Tyr Asn Asp Met
    130                 135                 140

Ile Thr Gln His Val Gln Arg His Lys Gly Ser Cys Gly Cys Arg Pro
145                 150                 155                 160

Leu Pro Thr Phe Gln Ile Asp Ala Phe Arg Arg Tyr Asp Ser Ser Pro
                165                 170                 175

Cys Ala Val Thr Cys Ser Asp Val Phe Gln Glu Cys Ser Tyr Asp Phe
            180                 185                 190

Gly Ser Gly Arg Asp Asn His Ala Val Ser Leu His Ser Ile Tyr Asp
        195                 200                 205

Ile Pro Tyr Ser Ser Ile Gly Pro Ala Leu His Arg Lys Asn Val Arg
    210                 215                 220

Val Cys Tyr Ala Ala Phe His Phe Ser Glu Ala Leu Leu Leu Gly Ser
225                 230                 235                 240

Pro Val Gly Asn Leu Asn Ser Ile Gly Ala Gln Phe Arg Val Asp Gly
                245                 250                 255

Asp Asp Val His Phe Leu Phe Ser Glu Glu Ser Thr Leu His Tyr Thr
            260                 265                 270

His Ser Leu Glu Asn Ile Lys Leu Ile Val Met Arg Thr Tyr Phe Pro
        275                 280                 285

Ala Asp Asp Arg Phe Val Tyr Ile Lys Glu Phe Met Val Lys Arg Val
    290                 295                 300
```

```
Asp Thr Phe Phe Phe Arg Leu Val Arg Ala Asp Thr His Met Leu His
305                 310                 315                 320

Lys Ser Val Gly His Tyr Ser Lys Ser Lys Ser Glu Tyr Phe Ala Leu
                325                 330                 335

Asn Thr Pro Pro Ile Phe Gln Asp Lys Ala Thr Phe Ser Val Trp Phe
            340                 345                 350

Pro Glu Ala Lys Lys Val Leu Ile Pro Lys Phe Glu Leu Ser Arg Phe
        355                 360                 365

Leu Ser Gly Asn Val Lys Ile Ser Arg Met Leu Val Asp Ala Asp Phe
    370                 375                 380

Val His Thr Ile Ile Asn His Ile Ser Thr Tyr Asp Asn Lys Ala Leu
385                 390                 395                 400

Val Trp Lys Asn Val Gln Ser Phe Val Glu Ser Ile Arg Ser Arg Val
                405                 410                 415

Ile Val Asn Gly Val Ser Val Lys Ser Glu Trp Asn Val Pro Val Asp
            420                 425                 430

Gln Leu Thr Asp Ile Ser Phe Ser Ile Phe Pro Leu Val Lys Val Arg
        435                 440                 445

Lys Val Gln Ile Glu Leu Met Ser Asp Lys Val Val Ile Glu Ala Arg
    450                 455                 460

Gly Leu Leu Arg Arg Phe Ala Asp Ser Leu Lys Ser Ala Val Glu Gly
465                 470                 475                 480

Leu Gly Asp Cys Val Tyr Asp Ala Leu Val Gln Thr Gly Trp Phe Asp
                485                 490                 495

Thr Ser Ser Asp Glu Leu Lys Val Leu Leu Pro Glu Pro Phe Met Thr
            500                 505                 510

Phe Ser Asp Tyr Leu Glu Gly Met Tyr Glu Ala Asp Ala Lys Ile Glu
        515                 520                 525

Arg Glu Ser Val Ser Glu Leu Leu Ala Ser Gly Asp Asp Leu Phe Lys
    530                 535                 540

Lys Ile Asp Glu Ile Arg Asn Asn Tyr Ser Gly Val Glu Phe Asp Val
545                 550                 555                 560

Glu Lys Phe Gln Glu Phe Cys Lys Glu Leu Asn Val Asn Pro Met Leu
                565                 570                 575

Ile Gly His Val Ile Glu Ala Ile Phe Ser Gln Lys Ala Gly Val Thr
            580                 585                 590

Val Thr Gly Leu Gly Thr Leu Ser Pro Glu Met Gly Ala Ser Val Ala
        595                 600                 605

Leu Ser Ser Thr Ser Val Asp Thr Cys Glu Asp Met Asp Val Thr Glu
    610                 615                 620

Asp Met Glu Asp Ile Val Leu Met Ala Asp Lys Ser His Ser Tyr Met
```

```
625                 630                 635                 640

Ser Pro Glu Met Ala Arg Trp Ala Asp Val Lys Tyr Gly Asn Asn Lys
                645                 650                 655

Gly Ala Leu Val Glu Tyr Lys Val Gly Thr Ser Met Thr Leu Pro Ala
            660                 665                 670

Thr Trp Ala Glu Lys Val Lys Ala Val Leu Pro Leu Ser Gly Ile Cys
        675                 680                 685

Val Arg Lys Pro Gln Phe Ser Lys Pro Leu Asp Glu Glu Asp Asp Leu
    690                 695                 700

Arg Leu Ser Asn Met Asn Phe Phe Lys Val Ser Asp Leu Lys Leu Lys
705                 710                 715                 720

Lys Thr Ile Thr Pro Val Val Tyr Thr Gly Thr Ile Arg Glu Arg Gln
                725                 730                 735

Met Lys Asn Tyr Ile Asp Tyr Leu Ser Ala Ser Leu Gly Ser Thr Leu
            740                 745                 750

Gly Asn Leu Glu Arg Ile Val Arg Ser Asp Trp Asn Gly Thr Glu Glu
        755                 760                 765

Ser Met Gln Thr Phe Gly Leu Tyr Asp Cys Glu Lys Cys Lys Trp Leu
    770                 775                 780

Leu Leu Pro Ala Glu Lys Lys His Ala Trp Ala Val Val Leu Ala Ser
785                 790                 795                 800

Asp Asp Thr Thr Arg Ile Ile Phe Leu Ser Tyr Asp Glu Ser Gly Ser
                805                 810                 815

Pro Ile Ile Asp Lys Lys Asn Trp Lys Arg Phe Ala Val Cys Ser Glu
            820                 825                 830

Thr Lys Val Tyr Ser Val Ile Arg Ser Leu Glu Val Leu Asn Lys Glu
        835                 840                 845

Ala Ile Val Asp Pro Gly Val His Ile Thr Leu Val Asp Gly Val Pro
    850                 855                 860

Gly Cys Gly Lys Thr Ala Glu Ile Ile Ala Arg Val Asn Trp Lys Thr
865                 870                 875                 880

Asp Leu Val Leu Thr Pro Gly Arg Glu Ala Ala Ala Met Ile Arg Arg
                885                 890                 895

Arg Ala Cys Ala Leu His Lys Ser Pro Val Ala Thr Asn Asp Asn Val
            900                 905                 910

Arg Thr Phe Asp Ser Phe Val Met Asn Arg Lys Ile Phe Lys Phe Asp
        915                 920                 925

Ala Val Tyr Val Asp Glu Gly Leu Met Val His Thr Gly Leu Leu Asn
    930                 935                 940

Phe Ala Leu Lys Ile Ser Gly Cys Lys Lys Ala Phe Val Phe Gly Asp
945                 950                 955                 960
```

```
Ala Lys Gln Ile Pro Phe Ile Asn Arg Val Met Asn Phe Asp Tyr Pro
                965                 970                 975

Lys Glu Leu Arg Thr Leu Ile Val Asp Asn Val Glu Arg Arg Tyr Val
            980                 985                 990

Thr His Arg Cys Pro Arg Asp Val Thr Ser Phe Leu Asn Thr Ile Tyr
        995                 1000                1005

Lys Ala Ala Val Ala Thr Thr Ser Pro Val Val His Ser Val Lys Ala
    1010                1015                1020

Ile Lys Val Ser Gly Ala Gly Ile Leu Arg Pro Glu Leu Thr Lys Ile
1025                1030                1035                1040

Lys Gly Lys Ile Ile Thr Phe Thr Gln Ser Asp Lys Gln Ser Leu Ile
                1045                1050                1055

Lys Ser Gly Tyr Asn Asp Val Asn Thr Val His Glu Ile Gln Gly Glu
            1060                1065                1070

Thr Phe Glu Glu Thr Ala Val Val Arg Ala Thr Pro Thr Pro Ile Gly
        1075                1080                1085

Leu Ile Ala Arg Asp Ser Pro His Val Leu Val Ala Leu Thr Arg His
    1090                1095                1100

Thr Lys Ala Met Val Tyr Tyr Thr Val Val Phe Asp Ala Val Thr Ser
1105                1110                1115                1120

Ile Ile Ala Asp Val Glu Lys Val Asp Gln Ser Ile Leu Thr Met Phe
                1125                1130                1135

Ala Thr Thr Val Pro Thr Lys
            1140


<210> 19
<211> 1503
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide
      sequence corresponding to 57 kD protein of CGMMV
      strain SH

<400> 19
atgcagaatt cgctgtatgt ccatcgtaat attttcctcc ctgttagtaa aacggggttt 60
tatacagaca tgcaggagtt ctacgataga tgccttcctg ggaattcctt cgtactaaat 120
gatttcgatg ccgtaaccat gcggttgagg gacaacgaat ttaacttaca accttgtagg 180
ctaaccttga gtaatttaga tccggtaccc gctttgatta agaatgaagc gcagaatttt 240
ctgatccccg ttttgcgtac ggcctgtgaa aggccgcgca ttccgggtct tcttgagaat 300
cttgtagcta tgataaagag gaatatgaat actcctgatt tagctgggac cgtagatata 360
actaacatgt cgatttctat agtagataac ttcttttctt cttttgttag ggacgaggtt 420
ttgcttgatc acttagattg tgttagggct agttccattc aaagtttttc tgattggttt 480
tcgtgtcaac caacctcagc ggttggccag ttagctaatt tcaatttcat agatttgcct 540
gcctttgata cttatatgca tatgattaag aggcaaccca agagtcggtt agatacttcg 600
attcagtctg aatatccggc cttgcaaact attgtttatc accctaaagt ggtaaatgca 660
```

```
gtttttggtc cggttttcaa gtatttaacc accaagtttc ttagtatggt agatagttct 720
aagtttttct tttacactag gaaaaaacca gaagatctgc aggaattttt ctcagatctc 780
tcttcccatt ctgattatga gattcttgag cttgatgttt ctaaatatga caagtcgcaa 840
tccgatttcc acttctctat tgagatggca atttgggaaa aattagggct tgacgatatt 900
ttggcttgga tgtggtctat gggtcacaaa agaactatac tgcaagattt ccaagccggg 960
ataaagacgc tcatttacta tcaacggaag tctggtgatg taactacttt tataggtaat 1020
acctttatta tcgcagcgtg tgtggctagt atgttgccgt tagataagtg ttttaaagct 1080
agtttttgtg gtgatgattc gctgatctac cttcctaagg gtttggagta tcctgatata 1140
caggctactg ccaacctigt ttggaatttt gaggcgaaac ttttccgaaa gaagtatggt 1200
tacttctgcg ggaagtatat aattcaccat gccaacggct gtattgttta ccctgaccct 1260
ttaaaattaa ttagtaaatt aggtaataag agtcttgtag ggtatgagca tgttgaggag 1320
tttcgtatat ctctcctcga cgttgctcat agtttgttta atggtgctta tttccattta 1380
ctcgacgatg caatccacga attatttcct aatgctgggg gttgcagttt tgtaattaat 1440
tgtttgtgta agtatttgag tgataagcgc cttttccgta gtctttacat agatgtctct 1500
aag                                                               1503
```

<210> 20
<211> 501
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: amino acid sequence corresponding to 57 kD protein of CGMMV strain SH

<400> 20

```
Met Gln Asn Ser Leu Tyr Val His Arg Asn Ile Phe Leu Pro Val Ser
 1               5                  10                  15

Lys Thr Gly Phe Tyr Thr Asp Met Gln Glu Phe Tyr Asp Arg Cys Leu
            20                  25                  30

Pro Gly Asn Ser Phe Val Leu Asn Asp Phe Asp Ala Val Thr Met Arg
        35                  40                  45

Leu Arg Asp Asn Glu Phe Asn Leu Gln Pro Cys Arg Leu Thr Leu Ser
    50                  55                  60

Asn Leu Asp Pro Val Pro Ala Leu Ile Lys Asn Glu Ala Gln Asn Phe
65                  70                  75                  80

Leu Ile Pro Val Leu Arg Thr Ala Cys Glu Arg Pro Arg Ile Pro Gly
                85                  90                  95

Leu Leu Glu Asn Leu Val Ala Met Ile Lys Arg Asn Met Asn Thr Pro
            100                 105                 110

Asp Leu Ala Gly Thr Val Asp Ile Thr Asn Met Ser Ile Ser Ile Val
        115                 120                 125

Asp Asn Phe Phe Ser Ser Phe Val Arg Asp Glu Val Leu Leu Asp His
    130                 135                 140

Leu Asp Cys Val Arg Ala Ser Ser Ile Gln Ser Phe Ser Asp Trp Phe
145                 150                 155                 160

Ser Cys Gln Pro Thr Ser Ala Val Gly Gln Leu Ala Asn Phe Asn Phe
                165                 170                 175
```

```
Ile Asp Leu Pro Ala Phe Asp Thr Tyr Met His Met Ile Lys Arg Gln
            180                 185                 190

Pro Lys Ser Arg Leu Asp Thr Ser Ile Gln Ser Glu Tyr Pro Ala Leu
        195                 200                 205

Gln Thr Ile Val Tyr His Pro Lys Val Val Asn Ala Val Phe Gly Pro
    210                 215                 220

Val Phe Lys Tyr Leu Thr Thr Lys Phe Leu Ser Met Val Asp Ser Ser
225                 230                 235                 240

Lys Phe Phe Phe Tyr Thr Arg Lys Lys Pro Glu Asp Leu Gln Glu Phe
                245                 250                 255

Phe Ser Asp Leu Ser Ser His Ser Asp Tyr Glu Ile Leu Glu Leu Asp
            260                 265                 270

Val Ser Lys Tyr Asp Lys Ser Gln Ser Asp Phe His Phe Ser Ile Glu
        275                 280                 285

Met Ala Ile Trp Glu Lys Leu Gly Leu Asp Asp Ile Leu Ala Trp Met
    290                 295                 300

Trp Ser Met Gly His Lys Arg Thr Ile Leu Gln Asp Phe Gln Ala Gly
305                 310                 315                 320

Ile Lys Thr Leu Ile Tyr Tyr Gln Arg Lys Ser Gly Asp Val Thr Thr
                325                 330                 335

Phe Ile Gly Asn Thr Phe Ile Ile Ala Ala Cys Val Ala Ser Met Leu
            340                 345                 350

Pro Leu Asp Lys Cys Phe Lys Ala Ser Phe Cys Gly Asp Asp Ser Leu
        355                 360                 365

Ile Tyr Leu Pro Lys Gly Leu Glu Tyr Pro Asp Ile Gln Ala Thr Ala
    370                 375                 380

Asn Leu Val Trp Asn Phe Glu Ala Lys Leu Phe Arg Lys Lys Tyr Gly
385                 390                 395                 400

Tyr Phe Cys Gly Lys Tyr Ile Ile His His Ala Asn Gly Cys Ile Val
                405                 410                 415

Tyr Pro Asp Pro Leu Lys Leu Ile Ser Lys Leu Gly Asn Lys Ser Leu
            420                 425                 430

Val Gly Tyr Glu His Val Glu Glu Phe Arg Ile Ser Leu Leu Asp Val
        435                 440                 445

Ala His Ser Leu Phe Asn Gly Ala Tyr Phe His Leu Leu Asp Asp Ala
    450                 455                 460

Ile His Glu Leu Phe Pro Asn Ala Gly Gly Cys Ser Phe Val Ile Asn
465                 470                 475                 480

Cys Leu Cys Lys Tyr Leu Ser Asp Lys Arg Leu Phe Arg Ser Leu Tyr
                485                 490                 495

Ile Asp Val Ser Lys
```

500

<210> 21
<211> 4932
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence corresponding to 186 kD protein of CGMMV strain SH

<400> 21
atggcaaaca ttaatgaaca aatcaacaac caacgtgacg ccgcggctag cgggagaaac 60
aatctcgtta gccaattggc gtcaaaaagg gtgtatgacg aggctgttcg ctcgttggat 120
catcaagaca gacgcccgaa aatgaatttt tctcgtgtgg tcagcacaga gcacaccagg 180
cttgtaactg acgcgtatcc ggagttttcg attagcttta ccgccaccaa gaactctgta 240
cactcccttg cgggtggtct gaggcttctt gaattggaat atatgatgat gcaggtgccc 300
tacggctcac cttgttatga catcggcggt aactatacgc agcacttgtt caaaggtaga 360
tcatatgtgc attgctgcaa tccgtgccta gatcttaagg atgttgcgag gaatgtgatg 420
tacaacgata tgattacgca acatgtacag aggcacaagg gatcttgcgg gtgcagacct 480
cttccaactt tccagataga tgcattcagg aggtacgata gttctccctg tgcggtcacc 540
tgttcagacg ttttccaaga gtgttcctat gattttggga gtggtaggga taatcatgca 600
gtctcgttgc attcaatcta cgatatccct tattcttcga tcggacctgc tcttcatagg 660
aagaatgtgc gagtttgtta tgcagccttt catttctcgg aggcattgct tttaggttcg 720
cctgtaggta atttaaatag tattggggct cagtttaggg tcgatggtga tgatgtgcat 780
tttcttttta gtgaagagtc tactttgcat tatactcata gtttagaaaa tatcaagtta 840
atcgtgatgc gtacttactt tcctgctgat gataggtttg tatatattaa ggagttcatg 900
gttaagcgtg tggatacttt tttctttagg ttggtcagag cagatacaca catgcttcat 960
aaatctgtgg ggcactattc gaaatcgaag tctgagtact tcgcgctgaa taccсctccg 1020
atcttccaag ataaagccac gttttctgtg tggtttcctg aagcgaagaa ggtgttgata 1080
cccaagtttg aactttcgag attcctttct gggaatgtga aaatctctag gatgcttgtc 1140
gatgctgatt tcgtccatac cattattaat cacattagca cgtatgataa caaggcctta 1200
gtgtggaaga atgttcagtc ctttgtggaa tccatacgtt caagagtaat tgtaaacgga 1260
gtttccgtga aatctgagtg gaacgtaccg gttgatcagc tcactgatat ctcgttctcg 1320
atattccctc tcgtgaaggt taggaaggta cagatcgagt taatgtctga taaagttgta 1380
atcgaggcga ggggtttgct tcggaggttc gcagacagtc ttaaatctgc cgtagaagga 1440
ctaggtgatt gcgtctatga tgctctagtt caaaccggct ggtttgacac ctctagcgac 1500
gaactgaaag tattgctacc tgaaccgttt atgacctttt cggattatct tgaagggatg 1560
tacgaggcag atgcaaagat cgagagagag agtgtctctg agttgctcgc ttccggtgat 1620
gatttgttca agaaaatcga tgagataaga aacaattaca gtggagtcga atttgatgta 1680
gagaaattcc aagaattttg caaggaactg aatgttaatc ctatgctaat tggccatgtc 1740
atcgaagcta tttttcgca gaaggctggg gtaacagtaa cgggtctggg cacgctctct 1800
cctgagatgg gcgcttctgt tgcgttatcc agtacctctg tagatacatg tgaagatatg 1860
gatgtaactg aagatatgga ggatatagtg ttgatggcgg acaagagtca ttcttacatg 1920
tcccctgaaa tggcgagatg ggctgatgtt aaatatggca acaataaagg ggctctagtc 1980
gagtacaaag tcggaacctc gatgacttta cctgccacct gggcagagaa agttaaggct 2040
gtcttaccgt tgtcggggat ctgtgtgagg aaaccccaat tttcgaagcc gcttgatgag 2100
gaagatgact tgaggttatc aaacatgaat ttctttaagg tgagcgatct aaagttgaag 2160
aagactatca ctccagtcgt ttacactggg accattcgag agaggcaaat gaagaattat 2220
attgattact tatcggcctc tcttggttcc acgctgggta atctggagag aatcgtgcgg 2280
agtgattgga atggtactga ggagagtatg caaacgttcg ggttgtatga ctgcgaaaag 2340
tgcaagtggt tattgttgcc agccgagaag aagcacgcat gggccgtggt tctggcaagt 2400
gacgatacca ctcgcataat cttcctttca tatgacgaat ctggttctcc tataattgat 2460
aagaaaaact ggaagcgatt tgctgtctgt tccgagacca aagtctatag tgtaattcgt 2520
agcttagagg ttctaaataa ggaagcaata gtcgaccccg ggttcacat aacattagtt 2580
gacggagtgc cgggttgtgg aaagaccgcc gagattatag cgagggtcaa ttggaaaact 2640
gatctagtat tgactcccgg aagggaggca gctgctatga ttaggcggag agcctgcgcc 2700
ctgcacaagt cacctgtggc aaccaatgac aacgtcagaa ctttcgattc ttttgtgatg 2760

aataggaaaa tcttcaagtt tgacgctgtc tatgttgacg agggtctgat ggtccatacg 2820
ggattactta attttgcgtt aaagatctca ggttgtaaaa aagccttcgt ctttggtgat 2880
gctaagcaaa tcccgtttat aaacagagtc atgaatttcg attatcctaa ggagttaaga 2940
actttaatag tcgataatgt agagcgtagg tatgtcaccc ataggtgtcc tagagatgtc 3000
actagttttc ttaatactat ctataaagcc gctgtcgcta ctactagtcc ggttgtacat 3060
tctgtgaagg caattaaagt gtcaggggcc ggtattctga ggcctgagtt gacaaagatc 3120
aaaggaaaga taataacgtt tactcaatct gataagcagt ccttgatcaa gagtgggtac 3180
aatgatgtga atactgtgca tgaaattcag ggagaaacct ttgaggagac ggcagttgtg 3240
cgtgccaccc cgactccaat aggtttgatt gcccgtgatt caccacatgt actagtggcc 3300
ttaactaggc acactaaggc aatggtgtat tatactgttg tattcgatgc agttacaagt 3360
ataatagcgg atgtggaaaa ggtcgatcag tcgatcttga ccatgtttgc taccactgtg 3420
cctaccaaaa tgcagaattc gctgtatgtc catcgtaata ttttcctccc tgttagtaaa 3480
acggggtttt atacagacat gcaggagttc tacgatagat gccttcctgg gaattccttc 3540
gtactaaatg atttcgatgc cgtaaccatg cggttgaggg acaacgaatt taacttacaa 3600
ccttgtaggc taaccttgag taatttagat ccggtacccg ctttgattaa gaatgaagcg 3660
cagaattttc tgatccccgt tttgcgtacg gcctgtgaaa ggccgcgcat tccgggtctt 3720
cttgagaatc ttgtagctat gataaagagg aatatgaata ctcctgattt agctgggacc 3780
gtagatataa ctaacatgtc gatttctata gtagataact tcttttcttc ttttgttagg 3840
gacgaggttt tgcttgatca cttagattgt gttagggcta gttccattca aagtttttct 3900
gattggtttt cgtgtcaacc aacctcagcg gttggccagt tagctaattt caatttcata 3960
gatttgcctg cctttgatac ttatatgcat atgattaaga ggcaacccaa gagtcggtta 4020
gatacttcga ttcagtctga atatccggcc ttgcaaacta ttgtttatca ccctaaagtg 4080
gtaaatgcag tttttggtcc ggttttcaag tatttaacca ccaagtttct tagtatggta 4140
gatagttcta agtttttctt ttacactagg aaaaaaccag aagatctgca ggaatttttc 4200
tcagatctct cttcccattc tgattatgag attcttgagc ttgatgtttc taaatatgac 4260
aagtcgcaat ccgatttcca cttctctatt gagatggcaa tttgggaaaa attagggctt 4320
gacgatattt tggcttggat gtggtctatg ggtcacaaaa gaactatact gcaagatttc 4380
caagccggga taaagacgct catttactat caacggaagt ctggtgatgt aactactttt 4440
ataggtaata cctttattat cgcagcgtgt gtggctagta tgttgccgtt agataagtgt 4500
tttaaagcta gtttttgtgg tgatgattcg ctgatctacc ttcctaaggg tttggagtat 4560
cctgatatac aggctactgc caaccttgtt tggaattttg aggcgaaact tttccgaaag 4620
aagtatggtt acttctgcgg gaagtatata attcaccatg ccaacggctg tattgtttac 4680
cctgaccctt taaaattaat tagtaaatta ggtaataaga gtcttgtagg gtatgagcat 4740
gttgaggagt ttcgtatatc tctcctcgac gttgctcata gtttgtttaa tggtgcttat 4800
ttccatttac tcgacgatgc aatccacgaa ttatttccta atgctggggg ttgcagtttt 4860
gtaattaatt gtttgtgtaa gtatttgagt gataagcgcc ttttccgtag tctttacata 4920
gatgtctcta ag 4932

<210> 22
<211> 1644
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: amino acid sequence corresponding to 186 kD protein of CGMMV strain SH

<400> 22
Met Ala Asn Ile Asn Glu Gln Ile Asn Asn Gln Arg Asp Ala Ala Ala
1 5 10 15

Ser Gly Arg Asn Asn Leu Val Ser Gln Leu Ala Ser Lys Arg Val Tyr
20 25 30

Asp Glu Ala Val Arg Ser Leu Asp His Gln Asp Arg Arg Pro Lys Met
35 40 45

Asn Phe Ser Arg Val Val Ser Thr Glu His Thr Arg Leu Val Thr Asp
50 55 60

```
Ala Tyr Pro Glu Phe Ser Ile Ser Phe Thr Ala Thr Lys Asn Ser Val
 65                 70                  75                  80

His Ser Leu Ala Gly Gly Leu Arg Leu Leu Glu Leu Glu Tyr Met Met
                85                  90                  95

Met Gln Val Pro Tyr Gly Ser Pro Cys Tyr Asp Ile Gly Gly Asn Tyr
            100                 105                 110

Thr Gln His Leu Phe Lys Gly Arg Ser Tyr Val His Cys Cys Asn Pro
        115                 120                 125

Cys Leu Asp Leu Lys Asp Val Ala Arg Asn Val Met Tyr Asn Asp Met
    130                 135                 140

Ile Thr Gln His Val Gln Arg His Lys Gly Ser Cys Gly Cys Arg Pro
145                 150                 155                 160

Leu Pro Thr Phe Gln Ile Asp Ala Phe Arg Arg Tyr Asp Ser Ser Pro
                165                 170                 175

Cys Ala Val Thr Cys Ser Asp Val Phe Gln Glu Cys Ser Tyr Asp Phe
            180                 185                 190

Gly Ser Gly Arg Asp Asn His Ala Val Ser Leu His Ser Ile Tyr Asp
        195                 200                 205

Ile Pro Tyr Ser Ser Ile Gly Pro Ala Leu His Arg Lys Asn Val Arg
    210                 215                 220

Val Cys Tyr Ala Ala Phe His Phe Ser Glu Ala Leu Leu Leu Gly Ser
225                 230                 235                 240

Pro Val Gly Asn Leu Asn Ser Ile Gly Ala Gln Phe Arg Val Asp Gly
                245                 250                 255

Asp Asp Val His Phe Leu Phe Ser Glu Glu Ser Thr Leu His Tyr Thr
            260                 265                 270

His Ser Leu Glu Asn Ile Lys Leu Ile Val Met Arg Thr Tyr Phe Pro
        275                 280                 285

Ala Asp Asp Arg Phe Val Tyr Ile Lys Glu Phe Met Val Lys Arg Val
    290                 295                 300

Asp Thr Phe Phe Phe Arg Leu Val Arg Ala Asp Thr His Met Leu His
305                 310                 315                 320

Lys Ser Val Gly His Tyr Ser Lys Ser Lys Ser Glu Tyr Phe Ala Leu
                325                 330                 335

Asn Thr Pro Pro Ile Phe Gln Asp Lys Ala Thr Phe Ser Val Trp Phe
            340                 345                 350

Pro Glu Ala Lys Lys Val Leu Ile Pro Lys Phe Glu Leu Ser Arg Phe
        355                 360                 365

Leu Ser Gly Asn Val Lys Ile Ser Arg Met Leu Val Asp Ala Asp Phe
    370                 375                 380
```

```
Val His Thr Ile Ile Asn His Ile Ser Thr Tyr Asp Asn Lys Ala Leu
385                 390                 395                 400

Val Trp Lys Asn Val Gln Ser Phe Val Glu Ser Ile Arg Ser Arg Val
                405                 410                 415

Ile Val Asn Gly Val Ser Val Lys Ser Glu Trp Asn Val Pro Val Asp
            420                 425                 430

Gln Leu Thr Asp Ile Ser Phe Ser Ile Phe Pro Leu Val Lys Val Arg
        435                 440                 445

Lys Val Gln Ile Glu Leu Met Ser Asp Lys Val Val Ile Glu Ala Arg
    450                 455                 460

Gly Leu Leu Arg Arg Phe Ala Asp Ser Leu Lys Ser Ala Val Glu Gly
465                 470                 475                 480

Leu Gly Asp Cys Val Tyr Asp Ala Leu Val Gln Thr Gly Trp Phe Asp
                485                 490                 495

Thr Ser Ser Asp Glu Leu Lys Val Leu Leu Pro Glu Pro Phe Met Thr
            500                 505                 510

Phe Ser Asp Tyr Leu Glu Gly Met Tyr Glu Ala Asp Ala Lys Ile Glu
        515                 520                 525

Arg Glu Ser Val Ser Glu Leu Leu Ala Ser Gly Asp Asp Leu Phe Lys
    530                 535                 540

Lys Ile Asp Glu Ile Arg Asn Asn Tyr Ser Gly Val Glu Phe Asp Val
545                 550                 555                 560

Glu Lys Phe Gln Glu Phe Cys Lys Glu Leu Asn Val Asn Pro Met Leu
                565                 570                 575

Ile Gly His Val Ile Glu Ala Ile Phe Ser Gln Lys Ala Gly Val Thr
            580                 585                 590

Val Thr Gly Leu Gly Thr Leu Ser Pro Glu Met Gly Ala Ser Val Ala
        595                 600                 605

Leu Ser Ser Thr Ser Val Asp Thr Cys Glu Asp Met Asp Val Thr Glu
    610                 615                 620

Asp Met Glu Asp Ile Val Leu Met Ala Asp Lys Ser His Ser Tyr Met
625                 630                 635                 640

Ser Pro Glu Met Ala Arg Trp Ala Asp Val Lys Tyr Gly Asn Asn Lys
                645                 650                 655

Gly Ala Leu Val Glu Tyr Lys Val Gly Thr Ser Met Thr Leu Pro Ala
            660                 665                 670

Thr Trp Ala Glu Lys Val Lys Ala Val Leu Pro Leu Ser Gly Ile Cys
        675                 680                 685

Val Arg Lys Pro Gln Phe Ser Lys Pro Leu Asp Glu Glu Asp Asp Leu
    690                 695                 700

Arg Leu Ser Asn Met Asn Phe Phe Lys Val Ser Asp Leu Lys Leu Lys
```

```
705                 710                 715                 720

Lys Thr Ile Thr Pro Val Val Tyr Thr Gly Thr Ile Arg Glu Arg Gln
                725                 730                 735

Met Lys Asn Tyr Ile Asp Tyr Leu Ser Ala Ser Leu Gly Ser Thr Leu
            740                 745                 750

Gly Asn Leu Glu Arg Ile Val Arg Ser Asp Trp Asn Gly Thr Glu Glu
        755                 760                 765

Ser Met Gln Thr Phe Gly Leu Tyr Asp Cys Glu Lys Cys Lys Trp Leu
    770                 775                 780

Leu Leu Pro Ala Glu Lys Lys His Ala Trp Ala Val Val Leu Ala Ser
785                 790                 795                 800

Asp Asp Thr Thr Arg Ile Ile Phe Leu Ser Tyr Asp Glu Ser Gly Ser
                805                 810                 815

Pro Ile Ile Asp Lys Lys Asn Trp Lys Arg Phe Ala Val Cys Ser Glu
            820                 825                 830

Thr Lys Val Tyr Ser Val Ile Arg Ser Leu Glu Val Leu Asn Lys Glu
        835                 840                 845

Ala Ile Val Asp Pro Gly Val His Ile Thr Leu Val Asp Gly Val Pro
    850                 855                 860

Gly Cys Gly Lys Thr Ala Glu Ile Ile Ala Arg Val Asn Trp Lys Thr
865                 870                 875                 880

Asp Leu Val Leu Thr Pro Gly Arg Glu Ala Ala Ala Met Ile Arg Arg
                885                 890                 895

Arg Ala Cys Ala Leu His Lys Ser Pro Val Ala Thr Asn Asp Asn Val
            900                 905                 910

Arg Thr Phe Asp Ser Phe Val Met Asn Arg Lys Ile Phe Lys Phe Asp
        915                 920                 925

Ala Val Tyr Val Asp Glu Gly Leu Met Val His Thr Gly Leu Leu Asn
    930                 935                 940

Phe Ala Leu Lys Ile Ser Gly Cys Lys Lys Ala Phe Val Phe Gly Asp
945                 950                 955                 960

Ala Lys Gln Ile Pro Phe Ile Asn Arg Val Met Asn Phe Asp Tyr Pro
                965                 970                 975

Lys Glu Leu Arg Thr Leu Ile Val Asp Asn Val Glu Arg Arg Tyr Val
            980                 985                 990

Thr His Arg Cys Pro Arg Asp Val Thr Ser Phe Leu Asn Thr Ile Tyr
        995                 1000                1005

Lys Ala Ala Val Ala Thr Thr Ser Pro Val Val His Ser Val Lys Ala
    1010                1015                1020

Ile Lys Val Ser Gly Ala Gly Ile Leu Arg Pro Glu Leu Thr Lys Ile
1025                1030                1035                1040
```

```
Lys Gly Lys Ile Ile Thr Phe Thr Gln Ser Asp Lys Gln Ser Leu Ile
                1045                1050                1055

Lys Ser Gly Tyr Asn Asp Val Asn Thr Val His Glu Ile Gln Gly Glu
            1060                1065                1070

Thr Phe Glu Glu Thr Ala Val Val Arg Ala Thr Pro Thr Pro Ile Gly
        1075                1080                1085

Leu Ile Ala Arg Asp Ser Pro His Val Leu Val Ala Leu Thr Arg His
    1090                1095                1100

Thr Lys Ala Met Val Tyr Tyr Thr Val Val Phe Asp Ala Val Thr Ser
1105                1110                1115                1120

Ile Ile Ala Asp Val Glu Lys Val Asp Gln Ser Ile Leu Thr Met Phe
                1125                1130                1135

Ala Thr Thr Val Pro Thr Lys Met Gln Asn Ser Leu Tyr Val His Arg
            1140                1145                1150

Asn Ile Phe Leu Pro Val Ser Lys Thr Gly Phe Tyr Thr Asp Met Gln
        1155                1160                1165

Glu Phe Tyr Asp Arg Cys Leu Pro Gly Asn Ser Phe Val Leu Asn Asp
    1170                1175                1180

Phe Asp Ala Val Thr Met Arg Leu Arg Asp Asn Glu Phe Asn Leu Gln
1185                1190                1195                1200

Pro Cys Arg Leu Thr Leu Ser Asn Leu Asp Pro Val Pro Ala Leu Ile
                1205                1210                1215

Lys Asn Glu Ala Gln Asn Phe Leu Ile Pro Val Leu Arg Thr Ala Cys
            1220                1225                1230

Glu Arg Pro Arg Ile Pro Gly Leu Leu Glu Asn Leu Val Ala Met Ile
        1235                1240                1245

Lys Arg Asn Met Asn Thr Pro Asp Leu Ala Gly Thr Val Asp Ile Thr
    1250                1255                1260

Asn Met Ser Ile Ser Ile Val Asp Asn Phe Phe Ser Ser Phe Val Arg
1265                1270                1275                1280

Asp Glu Val Leu Leu Asp His Leu Asp Cys Val Arg Ala Ser Ser Ile
                1285                1290                1295

Gln Ser Phe Ser Asp Trp Phe Ser Cys Gln Pro Thr Ser Ala Val Gly
            1300                1305                1310

Gln Leu Ala Asn Phe Asn Phe Ile Asp Leu Pro Ala Phe Asp Thr Tyr
        1315                1320                1325

Met His Met Ile Lys Arg Gln Pro Lys Ser Arg Leu Asp Thr Ser Ile
    1330                1335                1340

Gln Ser Glu Tyr Pro Ala Leu Gln Thr Ile Val Tyr His Pro Lys Val
1345                1350                1355                1360
```

```
Val Asn Ala Val Phe Gly Pro Val Phe Lys Tyr Leu Thr Thr Lys Phe
                1365                1370                1375

Leu Ser Met Val Asp Ser Ser Lys Phe Phe Phe Tyr Thr Arg Lys Lys
            1380                1385                1390

Pro Glu Asp Leu Gln Glu Phe Phe Ser Asp Leu Ser Ser His Ser Asp
        1395                1400                1405

Tyr Glu Ile Leu Glu Leu Asp Val Ser Lys Tyr Asp Lys Ser Gln Ser
    1410                1415                1420

Asp Phe His Phe Ser Ile Glu Met Ala Ile Trp Glu Lys Leu Gly Leu
1425                1430                1435                1440

Asp Asp Ile Leu Ala Trp Met Trp Ser Met Gly His Lys Arg Thr Ile
                1445                1450                1455

Leu Gln Asp Phe Gln Ala Gly Ile Lys Thr Leu Ile Tyr Tyr Gln Arg
            1460                1465                1470

Lys Ser Gly Asp Val Thr Thr Phe Ile Gly Asn Thr Phe Ile Ile Ala
        1475                1480                1485

Ala Cys Val Ala Ser Met Leu Pro Leu Asp Lys Cys Phe Lys Ala Ser
    1490                1495                1500

Phe Cys Gly Asp Asp Ser Leu Ile Tyr Leu Pro Lys Gly Leu Glu Tyr
1505                1510                1515                1520

Pro Asp Ile Gln Ala Thr Ala Asn Leu Val Trp Asn Phe Glu Ala Lys
                1525                1530                1535

Leu Phe Arg Lys Lys Tyr Gly Tyr Phe Cys Gly Lys Tyr Ile Ile His
            1540                1545                1550

His Ala Asn Gly Cys Ile Val Tyr Pro Asp Pro Leu Lys Leu Ile Ser
        1555                1560                1565

Lys Leu Gly Asn Lys Ser Leu Val Gly Tyr Glu His Val Glu Glu Phe
    1570                1575                1580

Arg Ile Ser Leu Leu Asp Val Ala His Ser Leu Phe Asn Gly Ala Tyr
1585                1590                1595                1600

Phe His Leu Leu Asp Asp Ala Ile His Glu Leu Phe Pro Asn Ala Gly
                1605                1610                1615

Gly Cys Ser Phe Val Ile Asn Cys Leu Cys Lys Tyr Leu Ser Asp Lys
            1620                1625                1630

Arg Leu Phe Arg Ser Leu Tyr Ile Asp Val Ser Lys
        1635                1640


<210> 23
<211> 1139
<212> DNA
<213> Artificial Sequence
```

```
<220>
<223> Description of Artificial Sequence: nucleotide
      sequence corresponding to sequence encoding coat
      protein of CGMMV strain SH

<400> 23
aattcggctt ctgtaggggt ggtgctactg ttgctctggt tgacacaagg atgcattctg 60
ttgcagaggg aactatatgc aaattttcag ctcccgccac cgtccgcgaa ttctctgtta 120
ggttcatacc taattatcct gtcgtggctg cggatgccct tcgcgatcct tggtctttat 180
ttgtgagact ctctaatgtg ggcattaaag atggtttcca tcctttgact ttagaggtcg 240
cttgtttagt cgctacaact aactctatta tcaaaaaggg tcttagagct tctgtagtcg 300
agtctgtcgt ctcttccgat cagtctattg tcctagattc cttgtccgag aaagttgaac 360
ctttctttga caaagttcct atttcagcgg ctgtaatggc aagagatccc agttataggt 420
ctaggtcaca gtctgtcggt ggtcgtggta agcggcattc taaacctcca aatcggaggt 480
tggactctgc ttctgaagag tccagttctg tttcttttga agatggctta caatccgatc 540
acacctagca aacttattgc gtttagtgct tcatatgttc ccgtcaggac tttacttaat 600
tttctagttg cttcacaagg taccgctttt cagactcaag cgggaagaga ttctttccgc 660
gagtccctgt ctgcgttacc ctcgtctgtc gtagatatta attctagatt cccagatgcg 720
ggtttttacg ctttcctcaa cggtcctgtg ttgaggccta tcttcgtttc gcttctcagc 780
tccacggata cgcgtaatag ggtcattgag gttgtagatc ctagcaatcc tacgactgct 840
gagtcgctta acgctgtaaa gcgtactgat gacgcgtcta cagccgctag ggccgagata 900
gataatttaa tagagtctat ttctaagggt tttgatgttt acgatagggc ttcatttgaa 960
gccgcgtttt cggtagtctg gtcagaggct accacctcga aagcttagtt tcgagggtct 1020
tctgatggtg gtgcacacca aagtgcatag tgctttcccg ttcacttaaa tcgaacggtt 1080
tgctcattgg tttgcggaaa cctctcacgt gtgacgttga agtttctatg ggcaagccg  1139


<210> 24
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 97G01

<400> 24
aggtgtcagt ggagaactca ttga                                        24


<210> 25
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 97G02

<400> 25
ggcgttgtgg tttgtgg                                                17


<210> 26
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 97G03

<400> 26
ctgtaggggt ggtgctactg t                                           21
```

<210> 27
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 97G18

<400> 27
gcccatagaa acttcaacgt c 21

<210> 28
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98A88

<400> 28
ccatggagaa ttcgctgtat gtcc 24

<210> 29
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98A86

<400> 29
cgagctctcg actgacacct tac 23

<210> 30
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98A84

<400> 30
ccatggcaaa cattaatgaa c 21

<210> 31
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer98A85

<400> 31
caaccatggc aaacattaat g 21

<210> 32
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98G63

<400> 32
taacagggag gaaaatatta cg 22

<210> 33
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L99

<400> 33
gagctcggat ccactagtaa cggc 24

<210> 34
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L107

<400> 34
tagagctctt gaagctaagc aaattccg 28

<210> 35
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L108

<400> 35
ttcaagagct ctaatcaccg aagacaaagg c 31

<210> 36
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L102

<400> 36
gaattatatc gattatctat cggc 24

<210> 37
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L103

<400> 37
gataatcgat ataattcttc atctgcc 27

<210> 38
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L104

<400> 38
aactagtaat tgatgatctg ttcaagaag 29

<210> 39
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L105

<400> 39
aattactagt ttccggaagc aagcagctca g 31

<210> 40
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer 98L106

<400> 40
gccctctaga tgcatgctcg ag 22

<210> 41
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
sequence 97G40

<400> 41
ctagagtttt aatttttata attaaacaaa 30

<210> 42
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
sequence 97G41

<400> 42
tcaaaattaa aaatattaat ttgtttgttg ttgttg 36

<210> 43
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
sequence 97G42

<400> 43
caacaacaac aacaacaaac aattttaaaa caacac 36

<210> 44
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
sequence 97G43

<400> 44
ttgttgtttg ttaaaatttt gttgtggtac 30